

XX WPI; 2000-246562/21.
 DR N-PSDB; AAZ51528.
 XX
 PT New human cell signaling proteins and polynucleotides useful for
 PT diagnosis, prevention and treatment of neoplastic, neurological,
 PT immunological, vesicle trafficking and smooth muscle disorders.
 XX
 PS Claim 1; Page 60; 62pp; English.
 XX
 CC The present sequence is human cell-signalling protein-2 (CSIG-2) involved
 CC in regulation of cell proliferation, differentiation and gene
 CC transcription. CSIG can be used in the diagnosis and treatment of
 CC diseases associated with expression of CSIG. These diseases include
 CC neoplastic, neurological, immunological, vesicle trafficking and smooth
 CC muscle disorders, including HIV, rheumatoid arthritis, asthma,
 CC atherosclerosis, diabetes mellitus, emphysema, irritable bowel syndrome,
 CC multiple sclerosis, osteoporosis, psoriasis and infections including
 CC viral, bacterial and fungal. CSIG and its catalytic or immunogenic
 CC fragments are useful for drug screening using libraries of compounds
 XX
 SQ Sequence 133 AA;
 Query Match 100.0%; Score 754; DB 3; Length 133;
 Best Local Similarity 100.0%; Pred. No. 1.9e-69;
 Matches 133; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 MAMVSAMSWLYLWISACAMLLCHGSLQHTFOQHLHLPREGGTCVIAAHRCCNKNRIE 60
 Db 1 MAMVSAMSWLYLWISACAMLLCHGSLQHTFOQHLHLPREGGTCVIAAHRCCNKNRIE 60
 QY 61 RSQTVKSCCLPGKVAGTTRNRPSCVDASIVIGKWCMEPCLEGECKTLPDMSGWMCAT 120
 Db 61 RSQTVKSCCLPGKVAGTTRNRPSCVDASIVIGKWCMEPCLEGECKTLPDMSGWMCAT 120
 QY 121 GNKIKTRIHPT 133
 Db 121 GNKIKTRIHPT 133
 RESULT 2
 AAM78667
 ID AAM78667 standard; protein; 133 AA.
 XX
 AC AAM78667;
 XX
 DT 06-NOV-2001 (first entry)
 XX
 DE Human protein SEQ ID NO 1329.
 XX
 KW Human; cytokine; cell proliferation; cell differentiation; gene therapy;
 KW vaccine; peptide therapy; stem cell growth factor; haematopoiesis;
 KW tissue growth factor; immunomodulatory; cancer; leukaemia;
 KW nervous system disorder; arthritis; inflammation.
 XX
 OS Homo sapiens.
 XX
 PN WO200157190-A2.
 XX
 PD 09-AUG-2001.
 XX
 PF 05-FEB-2001; 2001WO-US0004098.
 XX
 PR 03-FEB-2000; 2000US-00496914.
 PR 27-APR-2000; 2000US-00560875.
 PR 20-JUN-2000; 2000US-00598075.
 PR 19-JUL-2000; 2000US-00620325.
 PR 01-SEP-2000; 2000US-00654936.
 PR 15-SEP-2000; 2000US-00663561.
 PR 20-OCT-2000; 2000US-00693325.
 PR 30-NOV-2000; 2000US-00728422.
 XX
 PA (HYSE-) HYSEQ INC.

XX Tang YT, Liu C, Drmanac RT, Asundi V, Zhou P, Xu C, Cao Y;
 PI Ma Y, Zhao QA, Wang D, Wang J, Zhang J, Ren F, Chen R, Wang ZW;
 PI Xue AJ, Yang Y, Wejhrman T, Goodrich R;
 XX
 DR WPI; 2001-476283/51.
 DR N-PSDB; AAK51800.
 XX
 CC Nucleic acids encoding polypeptides with cytokine-like activities, useful
 CC in diagnosis and gene therapy.
 PT
 PT Claim 20; Page 3574; 6221pp; English.
 XX
 PS The invention relates to polynucleotides (AAK51456-AAK53435) and the
 CC encoded polypeptides (AAM78323-AAM80302) that exhibit activity elating to
 CC cytokine, cell proliferation or cell differentiation or which may induce
 CC production of other cytokines in other cell populations. The
 CC polynucleotides and polypeptides are useful in gene therapy, vaccines or
 CC peptide therapy. The polypeptides have various cytokine-like activities,
 CC e.g. stem cell growth factor activity, haematopoiesis regulating
 CC activity, tissue growth factor activity, immunomodulatory activity and
 CC activin/inhibin activity and may be useful in the diagnosis and/or
 CC treatment of cancer, leukaemia, nervous system disorders, arthritis and
 CC inflammation. Note: Records for SEQ ID NO 2110 (AAK52581), 2111
 CC (AAK52582) and 3686 (AAM80020) are omitted as the relevant pages from the
 CC sequence listing were missing at the time of publication
 XX
 SQ Sequence 133 AA;
 Query Match 100.0%; Score 754; DB 4; Length 133;
 Best Local Similarity 100.0%; Pred. No. 1.9e-69;
 Matches 133; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 MAMVSAMSWLYLWISACAMLLCHGSLQHTFOQHLHLPREGGTCVIAAHRCCNKNRIE 60
 Db 1 MAMVSAMSWLYLWISACAMLLCHGSLQHTFOQHLHLPREGGTCVIAAHRCCNKNRIE 60
 QY 61 RSQTVKSCCLPGKVAGTTRNRPSCVDASIVIGKWCMEPCLEGECKTLPDMSGWMCAT 120
 Db 61 RSQTVKSCCLPGKVAGTTRNRPSCVDASIVIGKWCMEPCLEGECKTLPDMSGWMCAT 120
 QY 121 GNKIKTRIHPT 133
 Db 121 GNKIKTRIHPT 133
 RESULT 3
 AAM79651
 ID AAM79651 standard; protein; 190 AA.
 XX
 AC AAM79651;
 XX
 DT 06-NOV-2001 (first entry)
 XX
 DE Human protein SEQ ID NO 3297.
 XX
 KW Human; cytokine; cell proliferation; cell differentiation; gene therapy;
 KW vaccine; peptide therapy; stem cell growth factor; haematopoiesis;
 KW tissue growth factor; immunomodulatory; cancer; leukaemia;
 KW nervous system disorder; arthritis; inflammation.
 XX
 OS Homo sapiens.
 XX
 PN WO200157190-A2.
 XX
 PD 09-AUG-2001.
 XX
 PF 05-FEB-2001; 2001WO-US0004098.
 XX
 PR 03-FEB-2000; 2000US-00496914.
 PR 27-APR-2000; 2000US-00560875.
 PR 20-JUN-2000; 2000US-00598075.
 PR 19-JUL-2000; 2000US-00620325.

PR 01-SEP-2000; 2000US-00654936.
 PR 15-SEP-2000; 2000US-00663561.
 PR 20-OCT-2000; 2000US-00693325.
 PR 30-NOV-2000; 2000US-00728422.

XX (HYSE-) HYSEQ INC.

PI Tang YT, Liu C, Drmanac RT, Asundi V, Zhou P, Xu C, Cao Y;
 PI Ma Y, Zhao QA, Wang D, Wang J, Zhang J, Ren F, Chen R, Wang ZW;
 PI Xue AJ, Yang Y, Wejhrman T, Goodrich R;

XX WPI: 2001-476283/51.

DR N-PSDB; AAK52784.

XX Nucleic acids encoding polypeptides with cytokine-like activities, useful
 PT in diagnosis and gene therapy.

XX Claim 20; Page 303; 6221pp; English.

CC The invention relates to polynucleotides (AAK51456-AAK53435) and the
 CC encoded polypeptides (AAM78323-AAM80302) that exhibit activity elating to
 CC cytokine, cell proliferation or cell differentiation or which may induce
 CC production of other cytokines in other cell populations. The
 CC polynucleotides and polypeptides are useful in gene therapy, vaccines or
 CC peptide therapy. The polypeptides have various cytokine-like activities,
 CC e.g. stem cell growth factor activity, haematopoiesis regulating
 CC activity, tissue growth factor activity, immunomodulatory activity and
 CC activin/inhibin activity and may be useful in the diagnosis and/or
 CC treatment of cancer, leukaemia, nervous system disorders, arthritis and
 CC inflammation. Note: Records for SEQ ID NO 2110 (AAK52581), 2111
 CC (AAK52582) and 3666 (AAM80020) are omitted as the relevant pages from the
 CC sequence listing were missing at the time of publication

XX Sequence 190 AA;

Query Match 100.0%; Score 754; DB 4; Length 190;
 Best Local Similarity 100.0%; Pred. No. 2.8e-69;
 Matches 133; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MAMVSAMSVLYLWISACAMLLCHGSLQHTFOQHLHRRPEGGTCEVIAAHRCCNKNRIE 60

DB 58 MAMVSAMSVLYLWISACAMLLCHGSLQHTFOQHLHRRPEGGTCEVIAAHRCCNKNRIE 117

QY 61 RSQTVKCSCLPGKVGATTRNRPSCVDASIVIGKWCMEPCLEGECKTLPDNSGWMCAT 120

DB 118 RSQTVKCSCLPGKVGATTRNRPSCVDASIVIGKWCMEPCLEGECKTLPDNSGWMCAT 177

QY 121 GNKIKTRIHPRT 133

DB 178 GNKIKTRIHPRT 190

RESULT 4

ABU52838

ID ABU52838 standard; protein; 133 AA.

XX AC ABU52838;

XX 14-APR-2003 (first entry)

XX Human kidney-derived protein from DKFZphfkd2_4b6.

XX Human; gene therapy; vaccine; disease treatment; detection.

XX Homo sapiens.

XX WO200112659-A2.

XX 22-FEB-2001.

XX 18-AUG-2000; 2000WO-IB001496.

XX 18-AUG-1999; 99US-0149499P.

PR 28-SEP-1999; 99US-0156503P.

XX (GEHU-) GERMAN HUMAN GENOME PROJECT.

XX Wiemann S;

XX WPI: 2001-327840/34.

DR N-PSDB; ABX71301.

XX Nucleic acids having the sequences of clones isolated from libraries of
 PT different human tissues, useful in recombinant DNA methodologies.

XX Claim 21; Page 426; 1095pp; English.

CC This invention describes novel polynucleotides and polypeptides isolated
 CC from human cDNA libraries which can be used for gene therapy or in
 CC vaccines. The polynucleotides of the invention and antibodies encoded by
 CC them may be used in the prevention, diagnosis and treatment of diseases
 CC associated with inappropriate polypeptide expression. The products of the
 CC invention may also be used to identify modulators of expression and
 CC activity and to down regulate expression and activity. The antibodies of
 CC the invention may also be used as diagnostic agents for detecting the
 CC presence of polypeptides in samples. This sequence represents a
 CC polypeptide described in the disclosure of the invention

XX Sequence 133 AA;

Query Match 98.9%; Score 746; DB 4; Length 133;

Best Local Similarity 99.2%; Pred. No. 1.3e-68;

Matches 132; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 MAMVSAMSVLYLWISACAMLLCHGSLQHTFOQHLHRRPEGGTCEVIAAHRCCNKNRIE 60

DB 1 MAMVSAMSVLYLWISACAMLLCHGSLQHTFOQHLHRRPEGGTCEVIAAHRCCNKNRIE 60

QY 61 RSQTVKCSCLPGKVGATTRNRPSCVDASIVIGKWCMEPCLEGECKTLPDNSGWMCAT 120

DB 61 RSQTVKCSCLPGKVGATTRNRPSCVDASIVIGKWCMEPCLEGECKTLPDNSGWMCAT 120

QY 121 GNKIKTRIHPRT 133

DB 121 GNKIKTRIHPRT 133

RESULT 5

AAM39191

ID AAM39191 standard; protein; 131 AA.

XX AC AAM39191;

XX 22-OCT-2001 (first entry)

XX Human polypeptide SEQ ID NO 2336.

XX Human; neotropic; immunosuppressant; cytostatic; gene therapy; cancer;
 KW peripheral nervous system; neuropathy; central nervous system; CNS;
 KW Alzheimer's, Parkinson's disease; Huntington's disease; haemostatic;
 KW amyotrophic lateral sclerosis; Shy-Drager Syndrome; chemotactic;
 KW chemokinetic; thrombolytic; drug screening; arthritis; inflammation;
 KW leukaemia.

XX Homo sapiens.

XX WO200153312-A1.

XX 26-JUL-2001.

XX 26-DEC-2000; 2000WO-US034263.

XX 23-DEC-1999; 99US-00471275.

XX 21-JAN-2000; 2000US-00488725.

XX 25-APR-2000; 2000US-0052317.

XX 20-JUN-2000; 2000US-00598042.

```

PR 19-JUL-2000; 2000US-00620312.
PR 03-AUG-2000; 2000US-00653450.
PR 14-SEP-2000; 2000US-00662191.
PR 19-OCT-2000; 2000US-00693036.
PR 29-NOV-2000; 2000US-00727344.
XX (HYSE-) HYSEQ INC.
PA
XX
XX Tang YT, Liu C, Asundi V, Chen R, Ma Y, Qian XB, Ren F, Wang D;
PI Wang J, Wang Z, Wehrman T, Xu C, Xue AJ, Yang Y, Zhang J, Zhao QA;
PI Zhou P, Goodrich R, Drmanac RT;
XX WPI; 2001-442253/47.
XX N-PSDB; AAI56347.
XX
XX Novel nucleic acids and polypeptides, useful for treating disorders such
PT as central nervous system injuries.
XX
XX Example 4; SEQ ID NO 2336; 10078pp; English.
XX
XX The invention relates to human nucleic acids (AAI57798-AAI61369) and the
CC encoded polypeptides (AAM38642-AAM42213) with nootropic,
CC immunosuppressant and cytostatic activity. The polynucleotides are useful
CC in gene therapy. A composition containing a polypeptide or polynucleotide
CC of the invention may be used to treat diseases of the peripheral nervous
CC system, such as peripheral nervous injuries, peripheral neuropathy and
CC localised neuropathies and central nervous system diseases, such as
CC Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic
CC lateral sclerosis, and Shy-Drager Syndrome. Other uses include the
CC utilisation of the activities such as: Immune system suppression,
CC Activin/inhibin activity, chemotactic/chemokinetic activity, haemostatic
CC and thrombolytic activity, cancer diagnosis and therapy, drug screening,
CC assays for receptor activity, arthritis and inflammation, leukaemias and
CC C.N.S disorders. Note: The sequence data for this patent did not form
CC part of the printed specification
XX
XX Sequence 131 AA;
XX
XX Query Match 57.3%; Score 432; DB 4; Length 131;
XX Best Local Similarity 61.2%; Pred. No. 2e-36;
XX Matches 74; Conservative 14; Mismatches 23; Indels 10; Gaps 1;
XX
QY 9 WVLYLWISACAMLLCHGSLQHTFQQHLLHRRPESGTCVIAAHRCCNKNRIERSQTVKCS 68
DB 19 FIVTLW-----GKVSSANHHKAAHVKTGTCEVVALHRCNKNRIERSQTVKCS 68
QY 69 CLPGKVAGTTRNRPSCVDASIVIGKWCMEBPCLEGECKTLPDNSGWMCATGNKIKTR 128
DB 69 CFFGQVAGTTTAAAPSCVDASIVEQKWCHMQPCLGECKVLPDRKGWSSGNGKVKTR 128
QY 129 I 129
DB 129 V 129
XX
XX RESULT 6
XX AAM40977
XX ID AAM40977 standard; protein; 242 AA.
XX AC AAM40977;
XX
XX 22-OCT-2001 (first entry)
XX
XX Human polypeptide SEQ ID NO 5908.
XX
XX Human; nootropic; immunosuppressant; cytostatic; gene therapy; cancer;
XX peripheral nervous system; neuropathy; central nervous system; CNS;
XX Alzheimer's; Parkinson's disease; Huntington's disease; haemostatic;
XX amyotrophic lateral sclerosis; Shy-Drager Syndrome; chemotactic;
XX chemokinetic; thrombolytic; drug screening; arthritis; inflammation;
XX leukaemia.
XX
XX Homo sapiens.
XX
XX Query Match 57.3%; Score 432; DB 4; Length 242;
XX Best Local Similarity 61.2%; Pred. No. 3.9e-36;
XX Matches 74; Conservative 14; Mismatches 23; Indels 10; Gaps 1;
XX
QY 9 WVLYLWISACAMLLCHGSLQHTFQQHLLHRRPESGTCVIAAHRCCNKNRIERSQTVKCS 68
DB 130 FIVTLW-----GKVSSANHHKAAHVKTGTCEVVALHRCNKNRIERSQTVKCS 179
QY 69 CLPGKVAGTTRNRPSCVDASIVIGKWCMEBPCLEGECKTLPDNSGWMCATGNKIKTR 128
DB 180 CFFGQVAGTTTAAAPSCVDASIVEQKWCHMQPCLGECKVLPDRKGWSSGNGKVKTR 239
QY 129 I 129
DB 240 V 240
XX
XX RESULT 7
XX AAB44958
XX ID AAB44958 standard; protein; 121 AA.
XX AC AAB44958;
XX
XX 12-FEB-2001 (first entry)
XX

```

XX DE Human secreted protein encoded by gene 42.

XX KW Secreted protein; human; immunosuppressive; antiarthritic; antirheumatic; antiproliferative; cytostatic; cardiant; vasotropic; cerebroprotective; neurotropic; neuroprotective; antibacterial; virucide; fungicide; ophthalmological; gene therapy; treatment; autoimmune disease; infection; cardiovascular disorder; nervous system disorder; ocular disorder; wound healing; epithelial cell proliferation; skin aging; mental state; transplantation; metabolism modulation.

XX OS Homo sapiens.

XX PN WO200005200-A1.

XX PD 21-SEP-2000.

XX PF 09-MAR-2000; 2000WO-US006042.

XX PR 12-MAR-1999; 99US-0124143P.

XX PR 03-DEC-1999; 99US-0168663P.

XX PA (HUMA-) HUMAN GENOME SCI INC.

XX PI Rosen CA, Ruben SM, Komatsoulis G;

XX DR WPI; 2000-656008/63.

XX DR N-PSDB; AAC79939.

XX PT Isolated human secretory proteins, nucleic acids encoding them and antibodies directed against them, useful for diagnosing and treating disorders related to the proteins such as cancer, Alzheimer's disease and Parkinsons.

XX PS Claim 11; Page 396-397; 453pp; English.

XX CC This invention describes a novel isolated polypeptide (I) and its encoding nucleic acid molecule (II) which have immunosuppressive, antiarthritic, antirheumatic, antiproliferative, cytostatic, cardiant, vasotropic, cerebroprotective, neurotropic, neuroprotective, antibacterial, virucide, fungicide and ophthalmological activity and which can be used for gene therapy. (I) and (II) are used to prevent, treat or ameliorate a medical condition in e.g. humans, mice, rabbits, goats, horses, cats, dogs, chickens or sheep. (I) and (II) are also used in diagnosing a pathological condition or susceptibility to a pathological condition. The antibodies to (II) can also be used in alleviating symptoms associated with the disorders and in diagnostic immunoassays e.g. radioimmunoassays or enzyme linked immunosorbent assays (ELISA). Disorders which are diagnosed or treated include autoimmune diseases e.g. rheumatoid arthritis, hyperproliferative disorders e.g. neoplasms of the breast or liver, cardiovascular disorders e.g. cardiac arrest, cerebrovascular disorders e.g. cerebral ischemia, angiogenesis, nervous system disorders e.g. Alzheimer's disease, infections caused by bacteria, viruses and fungi and ocular disorders e.g. corneal infection. The polypeptides can also be used to aid wound healing and epithelial cell proliferation, to prevent skin aging due to sunburn, to maintain organs before transplantation, for supporting cell culture of primary tissues, to regenerate tissues and in chemotaxis. The polypeptides are used to modulate mammalian metabolism, to change mammal's mental state or physical state by influencing biorhythms circadian rhythms, depression tendency for violence tolerance for pain, reproductive capabilities, hormonal or endocrine levels, appetite, libido, memory, stress or other cognitive qualities, as a food additive or preservative, such as to increase or decrease storage capabilities, fat content, lipid, protein, carbohydrates, vitamins, minerals, cofactors or other nutritional components

XX SQ Sequence 121 AA;

Query Match 32.0%; Score 241; DB 3; Length 121;

Best Local Similarity 38.4%; Pred. No. 7.2e-17;

Matches 48; Conservative 23; Mismatches 42; Indels 12; Gaps 2;

QY 6 AMSWVLYLWISACAMLLCHGSLQHTFOQHILHRPBGTCVIAAHRCCNKNRIERSQTV 65

DB 6 SMSSTFWAFMILASLLIAYCS-----QLAAGTCEIVTLDRSSQPRRTIARQTA 54

QY 66 KCSCLPGKVGAGTTRNPSQVDASIVIGKWCMEPCLEGECKTLPDMSGWMCATGSKI 124

DB 55 RCACRKGQAGTTRARPACVDARIITKTQWCDMLPCEGEGCDLLINRSWGTCTOPGGRI 114

QY 125 KTTIR 129

DB 115 KTTIV 119

RESULT 8

RAY53016

ID RAY53016 standard; protein; 132 AA.

XX AC RAY53016;

XX DT 29-FEB-2000 (first entry)

XX DE Human secreted protein clone ml243_1 protein sequence SEQ ID NO:38.

XX KW Human; secreted protein; nutritional; cytokine; cell proliferation; differentiation; immune stimulating; vaccine; suppression; haematopoiesis regulation; tissue growth; activin; inhibin; chemotactic; chemokinetic; haemostatic; thrombolytic; receptor; ligand; anti-inflammatory; cadherin; tumour invasion suppressor; tumour inhibition; gene therapy.

XX OS Homo sapiens.

XX PN WO9957132-A1.

XX PD 11-NOV-1999.

XX PF 07-MAY-1999; 99WO-US0009970.

XX PR 07-MAY-1998; 98US-0084564P.

XX PR 02-JUN-1998; 98US-0087645P.

XX PR 22-JUL-1998; 98US-0093712P.

XX PR 31-JUN-1998; 98US-0094935P.

XX PR 10-AUG-1998; 98US-0095880P.

XX PR 11-AUG-1998; 98US-0096068P.

XX PR 06-MAY-1999; 99US-00306111.

XX (GEMY) GENETICS INST INC.

XX PI Jacobs K, McCoy JM, Lavallie ER, Collins-Racie LA, Evans C; Merberg D, Treacy M, Agostino MJ, Steininger RJ, Bowman MR; Dibilasio-Smith E, Widom A;

XX DR WPI: 2000-052937/04.

XX DR N-PSDB; RAZ33334.

XX PT New polynucleotides encoding secreted human proteins, derived from adult placenta, adult retina, fetal brain, fetal.

XX PS Claim 47; Page 397; 492pp; English.

XX CC The present invention describes new human secreted proteins which were isolated from adult placenta, adult retina, foetal brain, foetal kidney, adult blood, adult brain, adult thyroid, adult bladder, adult neural tissue, adult testes, and adult lymph node CDNA libraries. The human secreted proteins, and the polynucleotides encoding them, are predicted to have biological activities which would make them suitable for treating, preventing or ameliorating medical conditions in humans and animals. Suggested activities include nutritional activity, cytokine and cell proliferation/differentiation activity, immune stimulating (e.g. as vaccine) or suppressing activity, haematopoiesis regulating activity, tissue growth activity, activin/inhibin activity, chemotactic/chemokinetic activity, haemostatic and thrombolytic activity, receptor/ligand activity, anti-inflammatory activity, cadherin/tumour

Query Match 31.9%; Score 240.5; DB 4; Length 89;
Best Local Similarity 49.4%; Pred. No. 5.7e-17;

QV 60 ERSOTVKCSCLPGKVAGTTTRNRPSCVD 86

Db 62 ERSQTVKSCFSGVAGTTAKPSCVD 88

RESULT 11

AAY13167

ID AAY13167 standard; protein; 29 AA.

XX AC AAY13167;

XX DT 22-JUN-1999 (first entry)

XX DE Human secreted protein encoded by 5' EST SEQ ID NO: 181.

XX KW Human; secreted protein; EST; expressed sequence tag; diagnosis;

XX KW forensic; gene therapy; chromosome mapping; signal peptide;

XX KW upstream regulatory sequence; cytokine activity; cell proliferation;

XX KW differentiation; haematopoiesis regulation; tissue growth regulation;

XX KW reproductive hormone regulation; chemotactic; chemokinetic; haemostatic;

XX KW thrombolytic; anti-inflammatory; tumour inhibition.

XX OS Homo sapiens.

XX AC WO9906552-A2.

XX PN 11-FEB-1999.

XX PD 31-JUL-1998; 98WO-IB001236.

XX PP 01-AUG-1997; 97US-00905223.

XX PR (GEST) GENSET.

XX PA Dumas Milne Edwards J, Duclert A, Lacroix B;

XX PI WPI; 1999-153782/13.

XX DR N-PSDB; AAX51967.

XX DX New isolated brain-derived nucleic acids - used to develop products which

XX PT may have cytokine, immune, regulatory, haematopoiesis regulating, anti-

XX PT inflammatory or tumour inhibition activity.

XX PS Claim 34; Page 548; 577pp; English.

XX CC AAX51787 to AAX52019 represent 5' expressed sequence tags (ESTs) for

XX CC human secreted proteins, and encode the proteins given in AAY12987 to

XX CC AAY13219, respectively. The proteins given represent the signal peptide

XX CC and an N-terminal fragment of a secreted protein. The nucleic acid

XX CC sequences can be used for producing secreted human gene products. They

XX CC can also be used to develop products for diagnosis and therapy. The

XX CC proteins obtained may have cytokine activity, cell

XX CC proliferation/differentiation activity, haematopoiesis regulating

XX CC activity, tissue growth regulating activity, reproductive hormone

XX CC thrombolytic activity, chemotactic/chemokinetic activity, haemostatic and

XX CC activity, tumour inhibition activity or other activities. The products

XX CC can be used in forensic, gene therapy and chromosome mapping procedures.

XX CC The sequences can also be used for obtaining corresponding promoter

XX CC sequences. The nucleic acids encoding the signal peptide can be used for

XX CC directing extracellular secretion of a polypeptide or the insertion of a

XX CC polypeptide into a membrane, or importing a polypeptide into a cell

XX CC Sequence 29 AA;

XX SQ

Query Match

Best Local Similarity 19.9%; Score 150; DB 2; Length 29;

Matches 28; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MAWVSAMSWLYLWISACAMLLCHGSLQ 28

Db 1 MAWVSAMSWLYLWISACAMLLCHGSLQ 28

RESULT 12

ADE60833

ID ADE60833 standard; protein; 1202 AA.

XX AC ADE60833;

XX DT 29-JAN-2004 (first entry)

XX DE Rat Protein P97607, SEQ ID NO 6745.

XX KW Rat; pain; neuronal tissue; gene therapy; spinal segmental nerve injury;

XX KW chronic constriction injury; CCI; spared nerve injury; SNI; Chung.

XX OS Rattus norvegicus.

XX PN WO2003016475-A2.

XX PD 27-FEB-2003.

XX PF 14-AUG-2002; 2002WO-US025765.

XX PR 14-AUG-2001; 2001US-0312147P.

XX PR 01-NOV-2001; 2001US-0346382P.

XX PR 26-NOV-2001; 2001US-0333347P.

XX XX (GEHO) GEN HOSPITAL CORP.

XX PA (FARB) BAYER AG.

XX XX Woolf C, D'urso D, Befort K, Costigan M;

XX PI WPI; 2003-268312/26.

XX DR GENBANK; P97607.

XX DX New composition comprising two or more isolated polypeptides, useful for

XX PT preparing a medicament for treating pain in an animal.

XX PS Claim 1; Page; 1017pp; English.

XX CC The invention discloses a composition comprising two or more isolated rat

XX CC or human polynucleotides or a polynucleotide which represents a fragment,

XX CC derivative or allelic variation of the nucleic acid sequence. Also

XX CC claimed are a vector comprising the novel polynucleotide, a host cell

XX CC comprising the vector, a method for identifying a nucleotide sequence

XX CC which is differentially regulated in an animal subjected to pain and a

XX CC kit to perform the method, an array, a method for identifying an agent

XX CC that increases or decreases the expression of the polynucleotide sequence

XX CC that is differentially expressed in neuronal tissue of a first animal

XX CC subjected to pain, a method for identifying a compound which regulates

XX CC the expression of a polynucleotide sequence which is differentially

XX CC expressed in an animal subjected to pain, a method for identifying a

XX CC compound that regulates the activity of one or more of the

XX CC polynucleotides, a method for producing a pharmaceutical composition, a

XX CC method for identifying a compound or small molecule that regulates the

XX CC activity in an animal of one or more of the polypeptides given in the

XX CC specification, a method for identifying a compound useful in treating

XX CC pain and a pharmaceutical composition comprising the one or more

XX CC polypeptides or their antibodies. The polynucleotide or the compound that

XX CC modulates its activity is useful for preparing a medicament for treating

XX CC pain (e.g. spinal segmental nerve injury (Chung), chronic constriction

XX CC injury (CCI) and spared nerve injury (SNI)) in an animal (e.g. gene

XX CC therapy). The sequence presented is a rat protein (shown in Table 2 of

XX CC the specification) which is differentially expressed during pain. Note:

XX CC The sequence data for this patent did not form part of the printed

XX CC specification, but was obtained in electronic form directly from WIPO at

XX CC ftp.wipo.int/pub/published_pct_sequences.

XX SQ Sequence 1202 AA;

Query Match

Best Local Similarity 14.7%; Score 110.5; DB 7; Length 1202;

Matches 33; Conservative 13; Mismatches 27; Indels 49; Gaps 9;

QY 23 CHGSLQHTFQQHLLHRRPEGTC-EVIAAHR-----CNKNRI-58

Db

RESULT 12

CC the expression of a polynucleotide sequence which is differentially
 CC expressed in an animal subjected to pain, a method for identifying a
 CC compound that regulates the activity of one or more of the
 CC polynucleotides, a method for producing a pharmaceutical composition, a
 CC method for identifying a compound or small molecule that regulates the
 CC activity in an animal of one or more of the polypeptides given in the
 CC specification, a method for identifying a compound useful in treating
 CC pain and a pharmaceutical composition comprising the one or more
 CC polypeptides or their antibodies. The polynucleotide or the compound that
 CC modulates its activity is useful for preparing a medicament for treating
 CC pain (e.g. spinal segmental nerve injury (Chung), chronic constriction
 CC injury (CCI) and spared nerve injury (SNI)) in an animal (e.g. gene
 CC therapy). The sequence presented is a rat protein (shown in Table 2 of
 CC the specification) which is differentially expressed during pain. Note:
 CC The sequence data for this patent did not form part of the printed
 CC specification, but was obtained in electronic form directly from WIPO at
 CC ftp.wipo.int/pub/published_pct_sequences.

XX Sequence 2531 AA;

Query Match 12.7%; Score 95.5; DB 7; Length 2531;
 Best Local Similarity 26.8%; Pred. No. 1.7;
 Matches 34; Conservative 7; Mismatches 35; Indels 51; Gaps 7;
 QY 15 ISACAMLLCHGSLQHTFQQHHLRPEGTCE-VIAAHC-----CNKN 56
 DB 679 IDECAGSPCH-----NGGTCEGDIAGFTCRCPGYHDPTCLSEVNECNSN 723
 QY 57 -----RIERSQTVKCSCLPGKVGATTRNRPSCVDASIVIGKWCMEPCLEGECKTLP 111
 DB 724 PCIHGACRDGLNGYKDCAPG-WSGTNCD-----INNNECESNPCVNGGTCKDM- 771
 QY 112 DNSGWM 118
 DB 772 -TSGYVC 777

RESULT 15

AD563705
 ID ADE63705 standard; protein; 2531 AA.
 AC ADE63705;
 XX 29-JAN-2004 (first entry)
 DT Rat Protein CAA40667, SEQ ID NO 9649.
 DE DE
 XX Rat; pain; neuronal tissue; gene therapy; spinal segmental nerve injury;
 KW chronic constriction injury; CCI; spared nerve injury; SNI; Chung.
 XX Rattus norvegicus.
 OS WO2003016475-A2.
 PN 27-FEB-2003.
 PD 14-AUG-2002; 2002WO-US025765.
 XX 14-AUG-2001; 2001US-0312147P.
 PR 01-NOV-2001; 2001US-0346382P.
 PR 26-NOV-2001; 2001US-0333347P.
 XX (GEO) GEN HOSPITAL CORP.
 PA (FARB) BAYER AG.
 XX Woolf C, D'urso D, Befort K, Costigan M;
 XX WEI; 2003-268312/26.
 DR GENBANK; CAA40667.
 XX New composition comprising two or more isolated polypeptides, useful for
 PT preparing a medicament for treating pain in an animal.
 XX

PS Claim 1; Page; 1017pp; English.
 XX The invention discloses a composition comprising two or more isolated rat
 CC or human polynucleotides or a polynucleotide which represents a fragment,
 CC derivative or allelic variation of the nucleic acid sequence. Also
 CC claimed are a vector comprising the novel polynucleotide, a host cell
 CC comprising the vector, a method for identifying a nucleotide sequence
 CC which is differentially regulated in an animal subjected to pain and a
 CC kit to perform the method, an array, a method for identifying an agent
 CC that increases or decreases the expression of the polynucleotide sequence
 CC that is differentially expressed in neuronal tissue of a first animal
 CC subjected to pain, a method for identifying a compound which regulates
 CC the expression of a polynucleotide sequence which is differentially
 CC expressed in an animal subjected to pain, a method for identifying a
 CC compound that regulates the activity of one or more of the
 CC polynucleotides, a method for producing a pharmaceutical composition, a
 CC method for identifying a compound or small molecule that regulates the
 CC activity in an animal of one or more of the polypeptides given in the
 CC specification, a method for identifying a compound useful in treating
 CC pain and a pharmaceutical composition comprising the one or more
 CC polypeptides or their antibodies. The polynucleotide or the compound that
 CC modulates its activity is useful for preparing a medicament for treating
 CC pain (e.g. spinal segmental nerve injury (Chung), chronic constriction
 CC injury (CCI) and spared nerve injury (SNI)) in an animal (e.g. gene
 CC therapy). The sequence presented is a rat protein (shown in Table 2 of
 CC the specification) which is differentially expressed during pain. Note:
 CC The sequence data for this patent did not form part of the printed
 CC specification, but was obtained in electronic form directly from WIPO at
 CC ftp.wipo.int/pub/published_pct_sequences.

XX Sequence 2531 AA;

Query Match 12.7%; Score 95.5; DB 7; Length 2531;
 Best Local Similarity 26.8%; Pred. No. 1.7;
 Matches 34; Conservative 7; Mismatches 35; Indels 51; Gaps 7;
 QY 15 ISACAMLLCHGSLQHTFQQHHLRPEGTCE-VIAAHC-----CNKN 56
 DB 679 IDECAGSPCH-----NGGTCEGDIAGFTCRCPGYHDPTCLSEVNECNSN 723
 QY 57 -----RIERSQTVKCSCLPGKVGATTRNRPSCVDASIVIGKWCMEPCLEGECKTLP 111
 DB 724 PCIHGACRDGLNGYKDCAPG-WSGTNCD-----INNNECESNPCVNGGTCKDM- 771
 QY 112 DNSGWM 118
 DB 772 -TSGYVC 777

Search completed: April 8, 2004, 13:16:11
 Job time : 58 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 8, 2004, 13:14:59 ; Search time 18 Seconds
(without alignments)
384.741 Million cell updates/sec

Title: US-09-763-335-2

Perfect score: 754

Sequence: 1 MAMVSNMSVLYLWISACAM.....SCHWMCATGNKIKXTRIHPRT 133

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 141681 seqs, 52070155 residues

Total number of hits satisfying chosen parameters: 141681

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt_42.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	110.5	14.7	1202	1	JAG2 RAT
2	96	12.7	1964	1	NTC4 MOUSE
3	95.5	12.7	2531	1	NTC1 MOUSE
4	95.5	12.7	2531	1	NTC1 RAT
5	92.5	12.3	2139	1	CRB_DROME
6	91.5	12.1	1213	1	JAG3 BRARE
7	91	12.1	555	1	DP87 DICDI
8	86.5	11.5	2703	1	NOTC DROME
9	86	11.4	685	1	DLA4 HUMAN
10	86	11.4	1218	1	JAG1 HUMAN
11	86	11.4	1218	1	JAG1 MOUSE
12	86	11.4	1219	1	JAG1 RAT
13	85.5	11.3	1064	1	FBP1 STRPU
14	84.5	11.2	2003	1	NTC4 HUMAN
15	84	11.1	1247	1	JAG2 MOUSE
16	84	11.1	2524	1	NOTC XENLA
17	83.5	11.1	1170	1	TSP2 BOVIN
18	82.5	10.9	2470	1	NTC2 MOUSE
19	82.5	10.9	2471	1	NTC2 HUMAN
20	82.5	10.9	2471	1	NTC2 RAT
21	82	10.9	2318	1	NTC3 MOUSE
22	82	10.9	4544	1	LRP1 HUMAN
23	82	10.9	5376	1	ZAN MOUSE
24	81.5	10.8	355	1	YMT9 YEAST
25	81.5	10.8	2556	1	NTC1 HUMAN
26	80.5	10.7	2321	1	NTC3 HUMAN
27	80	10.6	937	1	VWF BOVIN
28	79.5	10.5	2319	1	NTC3 RAT
29	79	10.5	819	1	AD09_HUMAN
30	78.5	10.4	1178	1	TSP2 CHICK
31	78.5	10.4	3312	1	CLR3 HUMAN
32	78	10.3	1191	1	LMG2 MOUSE
33	78	10.3	1376	1	CRBH_HUMAN

34	78	10.3	2311	1	FBM2 HUMAN
35	78	10.3	4543	1	LRP1 CHICK
36	77.5	10.3	686	1	DLA4 MOUSE
37	77.5	10.3	723	1	DLA1 HUMAN
38	77.5	10.3	1125	1	TIR2 BOVIN
39	77.5	10.3	1877	1	PKX5 MOUSE
40	77.5	10.3	2437	1	NTC1 BRARE
41	77	10.2	843	1	CO7 HUMAN
42	77	10.2	1403	1	NID2 MOUSE
43	77	10.2	2907	1	FBM2 MOUSE
44	76.5	10.1	1172	1	TSP2 MOUSE
45	76	10.1	592	1	DLA3 MOUSE

ALIGNMENTS

RESULT 1
JAG2 RAT
ID JAG2 RAT STANDARD; PRT; 1202 AA.
AC P97607;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 15-MAR-2004 (Rel. 43, Last annotation update)
DE Jagged 2 (Jagged2) (Fragment).
GN JAG2.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Rattus.
OX NCBI_TaxID:10116;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Brain;
RX MEDLINE=97105052; PubMed=8948600;
RA Shawber C., Boulter J., Lindseil C.E., Weinmaster G.;
RT "Jagged2: a serrate-like gene expressed during rat embryogenesis.";
RL Dev. Biol. 180:370-376(1996).
CC -!- FUNCTION: Putative Notch ligand involved in the mediation of Notch signaling. May have a role in neurogenesis in the peripheral nervous system, limb development and in the adult brain.
CC -!- SUBCELLULAR LOCATION: Type I membrane protein.
CC -!- DEVELOPMENTAL STAGE: At stage E12.5 it is detected in dorsal root ganglia, AER, and surface ectoderm. At E14.5, found as well in cranial ganglia, thymus and olfactory epithelia. At E16.5, found as well in salivary gland, tooth buds and hair follicles.
CC -!- SIMILARITY: Contains 16 EGF-like domains.
CC -!- SIMILARITY: Contains 1 DSL domain.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).
CC EMBL; U70050; AAC52946.1; -.
CC HSP; P00743; ICCF.
CC GO; GO:0005887; C:integral to plasma membrane; ISS.
CC GO; GO:0008083; P:growth factor activity; ISS.
CC GO; GO:0005112; P:Notch binding; NAS.
CC GO; GO:0007049; P:cell cycle; ISS.
CC GO; GO:0001054; P:cell differentiation; NAS.
CC GO; GO:0001709; P:cell fate determination; NAS.
CC GO; GO:0009912; P:hair cell fate commitment; ISS.
CC GO; GO:0007605; P:hearing; ISS.
CC GO; GO:0003026; P:limb morphogenesis; NAS.
CC GO; GO:0007219; P:N signaling pathway; NAS.
CC GO; GO:0003034; P:regulation of cell migration; ISS.
CC GO; GO:0042127; P:regulation of cell proliferation; ISS.
CC GO; GO:0007283; P:spermatogenesis; IEP.
CC GO; GO:0030217; P:T-cell differentiation; ISS.
CC GO; GO:0045061; P:thymic T-cell selection; ISS.

DR InterPro; IPR000152; Asx_hydroxyl_S.
 DR InterPro; IPR001774; DSL.
 DR InterPro; IPR000742; EGF_2.
 DR InterPro; IPR001881; EGF_Ca.
 DR InterPro; IPR001438; EGF_II.
 DR InterPro; IPR006209; EGF_Like.
 DR InterPro; IPR001007; VWF_C.
 DR Pfam; PF01414; DSL; 1.
 DR Pfam; PF00008; EGF; 14.
 DR PRINTS; PR00010; EGFBLD.
 DR SMART; SM00051; DSL; 1.
 DR SMART; SM00179; EGF_CA; 7.
 DR SMART; SM00214; VWF_1.
 DR PROSITE; PS00010; ASX_HYDROXYL; 10.
 DR PROSITE; PS00022; EGF_1; 15.
 DR PROSITE; PS01186; EGF_2; 11.
 DR PROSITE; PS00026; EGF_3; 15.
 DR PROSITE; PS01187; EGF_CA; 7.
 KW Calcium-binding; EGF-Like domain; Glycoprotein; Developmental protein;
 KW Repeat; Transmembrane.
 FT NON_TER 1 1
 FT DOMAIN <1 1085
 FT TRANSMEM 1086 1102
 FT DOMAIN 1103 1202
 FT DOMAIN 132 194
 FT DOMAIN 195 228
 FT DOMAIN 229 259
 FT DOMAIN 261 299
 FT DOMAIN 301 337
 FT DOMAIN 339 375
 FT DOMAIN 377 413
 FT DOMAIN 415 450
 FT DOMAIN 452 488
 FT DOMAIN 490 527
 FT DOMAIN 529 589
 FT DOMAIN 591 627
 FT DOMAIN 629 665
 FT DOMAIN 667 703
 FT DOMAIN 706 742
 FT DOMAIN 744 780
 FT DOMAIN 782 818
 FT DISULFID 199 210
 FT DISULFID 203 216
 FT DISULFID 218 227
 FT DISULFID 230 241
 FT DISULFID 236 247
 FT DISULFID 249 258
 FT DISULFID 265 277
 FT DISULFID 271 287
 FT DISULFID 289 298
 FT DISULFID 305 316
 FT DISULFID 310 325
 FT DISULFID 327 336
 FT DISULFID 343 354
 FT DISULFID 348 363
 FT DISULFID 365 374
 FT DISULFID 381 392
 FT DISULFID 386 401
 FT DISULFID 403 412
 FT DISULFID 419 429
 FT DISULFID 423 438
 FT DISULFID 440 449
 FT DISULFID 456 467
 FT DISULFID 461 476
 FT DISULFID 478 487
 FT DISULFID 495 506
 FT DISULFID 500 515
 FT DISULFID 517 526
 FT DISULFID 544 567
 FT DISULFID 561 577
 FT DISULFID 579 588
 FT DISULFID 595 606
 FT DISULFID 600 615

FT DISULFID 617 626 BY SIMILARITY.
 FT DISULFID 633 644 BY SIMILARITY.
 FT DISULFID 638 653 BY SIMILARITY.
 FT DISULFID 655 664 BY SIMILARITY.
 FT DISULFID 671 682 BY SIMILARITY.
 FT DISULFID 676 691 BY SIMILARITY.
 FT DISULFID 693 702 BY SIMILARITY.
 FT DISULFID 710 721 BY SIMILARITY.
 FT DISULFID 715 730 BY SIMILARITY.
 FT DISULFID 732 741 BY SIMILARITY.
 FT DISULFID 748 759 BY SIMILARITY.
 FT DISULFID 753 768 BY SIMILARITY.
 FT DISULFID 770 779 BY SIMILARITY.
 FT DISULFID 786 797 BY SIMILARITY.
 FT DISULFID 791 806 BY SIMILARITY.
 FT DISULFID 808 817 BY SIMILARITY.
 SQ SEQUENCE 1202 AA; 129703 MW; 08CB4E5271FF8BE CRC64;
 Query Match 14.7%; Score 110.5; DB 1; Length 1202;
 Best Local Similarity 27.0%; Pred. No. 0.0029;
 Matches 33; Conservative 13; Mismatches 27; Indels 49; Gaps 9;
 QY 23 CHGSLQHTFQOHLHRPEGGTC-EVIAAHC-----CNKNRI- 58
 DB 419 CHGQCOH-----GTCCKDLVNGVQCVRGFGGRHCELEYKCASSPCRRGGIC 467
 QY 59 ERSQTVKSCSLPGKVGAGTRNRPS-CVDASIVIGKWCMEPCLEGECCKTLDPNSGM 117
 DB 468 EDLVGDFRCHC-PRGLSG-----PLCEVDVDL-----WCEPNPCLNGARCYNLEDD--YY 514
 QY 118 CA 119
 DB 515 CA 516
 RESULT 2
 NTC4 MOUSE STANDARD; PRT; 1964 AA.
 ID NTC4 MOUSE STANDARD; PRT; 1964 AA.
 AC P31695; O35442; O88314; O62390; Q9R1W9; Q9R1X0;
 DT 01-JUL-1993 (Rel. 26, Created)
 DT 01-NOV-1997 (Rel. 35, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Neurogenic locus notch homolog protein 4 precursor (Notch 4)
 DE [Contains: Transforming protein Int-3].
 GN NOTCH4 OR INT3 OR INT-3.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=92194507; PubMed=1312643;
 RA Robbins J., Blondel B.J., Gallahan D., Callahan R.;
 RT "Mouse mammary tumor gene int-3: a member of the notch gene family
 RT transforms mammary epithelial cells.";
 RL J. Virol. 66:2594-2599(1992).
 RN [2]
 RP REVISIONS, SEQUENCE FROM N.A.
 RX MEDLINE=97294599; PubMed=9150355;
 RA Gallahan D., Callahan R.;
 RT "The mouse mammary tumor associated gene INT3 is a unique member of
 RT the NOTCH gene family (NOTCH4).";
 RL Oncogene 14:1893-1896(1997).
 RN [3]
 RP SEQUENCE FROM N.A.
 RX TISSUE=Lung, and Testis;
 RA MEDLINE=96281668; PubMed=8681805;
 RT Uyttendaele H., Marazzi G., Wu G., Yan Q., Sassoon D., Kitajewski J.;
 RT "Notch4/int-3, a mammary proto-oncogene, is an endothelial
 RT cell-specific mammalian Notch gene.";
 RL Development 122:2251-2259(1996).
 RN [4]
 RP SEQUENCE FROM N.A.

RA Rowen L., Mahairas G., Qin S., Ahearn M.E., Dankers C., Lasky S.,
 RA Loretz C., Schmidt S., Tipton S., Traicoff R., Zackrone K., Hood L.,
 RT "Sequence of the mouse major histocompatibility locus class III
 RL region.";
 RL Submitted (OCT-1997) to the EMBL/GenBank/DBJ databases.
 RN [5]
 RP SEQUENCE OF 1436-1600 FROM N.A.
 RX MEDLINE=99252212; PubMed=10233392;
 RA Lee J.-S., Haruna T., Ishimoto A., Honjo T., Yanagawa S.-I.,
 RT "Intracisternal type A particle-mediated activation of the Notch4/int3
 RT gene in a mouse mammary tumor: generation of truncated Notch4/int3
 RT mRNAs by retroviral splicing events.";
 RL J. Virol. 73:5166-5171 (1999).
 RN [6]
 RP FUNCTION.
 RX MEDLINE=21244657; PubMed=11344305;
 RA Uyttendaele H., Ho J., Rossant J., Kitajewski J.,
 RT "Vascular patterning defects associated with expression of activated
 RT Notch4 in embryonic endoderm.";
 RL Proc. Natl. Acad. Sci. U.S.A. 98:5643-5648 (2001).
 RN [7]
 RP SEQUENCE OF 1463-1964, POST-TRANSLATIONAL PROCESSING, AND MUTAGENESIS
 OF VAL-1463.
 RX MEDLINE=215233956; PubMed=11518718;
 RA Saxena M.I., Schroeter E.H., Mumm J.S., Kopan R.,
 RT "Murine notch homologs (N1-4) undergo presenilin-dependent
 RT proteolysis.";
 RL J. Biol. Chem. 276:40268-40273 (2001).
 RN [8]
 RP POST-TRANSLATIONAL PROCESSING.
 RX MEDLINE=21374376; PubMed=11459941;
 RA Mizutani T., Taniguchi Y., Aoki T., Hashimoto N., Honjo T.,
 RT "Conservation of the biochemical mechanisms of signal transduction
 RT among mammalian Notch family members.";
 RL Proc. Natl. Acad. Sci. U.S.A. 98:9026-9031 (2001).
 CC -!- FUNCTION: Functions as a receptor for membrane-bound ligands
 CC Jagged1, Jagged2 and Deltal to regulate cell-fate determination.
 CC Upon ligand activation through the released notch intracellular
 CC domain (NICD) it forms a transcriptional activator complex with
 CC RBP-J kappa and activates genes of the enhancer of split locus.
 CC Affects the implementation of differentiation, proliferation and
 CC apoptotic programs (By similarity). May regulate branching
 CC morphogenesis in the developing vascular system.
 CC -!- SUBUNIT: Heterodimer of a C-terminal fragment N(TM) and a N-
 CC terminal fragment N(EC) which are probably linked by disulfide
 CC bonds.
 CC -!- SUBCELLULAR LOCATION: Type I membrane protein. Following
 CC proteolytical processing NICD is translocated to the nucleus.
 CC -!- TISSUE SPECIFICITY: Highly expressed in lung, moderately in heart
 CC kidney, and at lower levels in the ovary and skeletal muscle. A
 CC very low expression is seen in the brain, intestine, liver and
 CC testis.
 CC -!- DEVELOPMENTAL STAGE: Highly expressed in endothelial cells during
 CC embryonic development from 9.0 dpc.
 CC -!- PTM: Synthesized in the endoplasmic reticulum as an inactive form
 CC which is proteolytically cleaved by a furin-like convertase in the
 CC trans-Golgi network before it reaches the plasma membrane to yield
 CC an active, ligand-accessible form. Cleavage results in a C-
 CC terminal fragment N(TM) and a N-terminal fragment N(EC). Following
 CC ligand binding, it is cleaved by TNF-alpha converting enzyme
 CC (TACE) to yield a membrane-associated intermediate fragment called
 CC notch extracellular truncation (NEXT). This fragment is then
 CC cleaved by presenilin dependent gamma-secretase to release a
 CC notch-derived peptide containing the intracellular domain (NICD)
 CC from the membrane.
 CC -!- PTM: Phosphorylated.
 CC -!- DISEASE: Loss of the extracellular domain causes constitutive
 CC activation of the Notch protein, which leads to hyperproliferation
 CC of glandular epithelial tissues and development of mammary
 CC carcinomas.
 CC -!- SIMILARITY: Belongs to the NOTCH family.
 CC -!- SIMILARITY: Contains 29 EGF-like domains.
 CC -!- SIMILARITY: Contains 3 Lin/Notch repeats.

CC -!- SIMILARITY: Contains 5 ANK repeats.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See [http://www.isb-sib.ch/](http://www.isb-sib.ch/announce/)
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL: M80456; AAB38377.1; -;
 CC EMBL: U43691; AAC52630.1; -;
 CC EMBL: U43691; AAC52631.1; -;
 CC EMBL: AF030001; AAB82004.1; -;
 CC EMBL: AB016771; BAA32281.1; ALT SEQ.
 CC EMBL: AB016772; BAA32283.1; ALT INIT.
 CC EMBL: AB016773; BAA32284.1; ALT_INIT.
 CC EMBL: AB016774; BAA32285.1; -;
 CC FIR: A38072; TWMV73.
 CC FIR: T09059; T09059.
 CC HSSP: P08709; IBF9.
 CC MGD: MGI:107471; Notch4.
 CC InterPro: IPR002110; ANK.
 CC InterPro: IPR000152; Asx_hydroxyl_S.
 CC InterPro: IPR000742; EGF_2.
 CC InterPro: IPR001881; EGF_Ca.
 CC InterPro: IPR001438; EGF_II.
 CC InterPro: IPR006209; EGF_like.
 CC InterPro: IPR002049; Laminin_EGF.
 CC InterPro: IPR008297; Notch.
 CC InterPro: IPR008000; Notch_dom.
 CC Pfam: PF00023; ank; 6.
 CC Pfam: PF00008; EGF; 27.
 CC Pfam: PF00066; notch; 2.
 CC PIRSF: PIRSF002279; Notch; 1.
 CC PRINTS: PR00010; EGFBL00D.
 CC PRINTS: PR00011; EGFFLAMININ.
 CC PRINTS: PR01452; NOTCH.
 CC SMART: SM00248; ANK; 6.
 CC SMART: SM00179; EGF_CA; 11.
 CC SMART: SM00004; NL; 2.
 CC PROSITE: PS00297; ANK_REPEAT; 1.
 CC PROSITE: PS00088; ANK_REPEAT; 5.
 CC PROSITE: PS00010; ASX_HYDROXYL; 11.
 CC PROSITE: PS00022; EGF_1; 28.
 CC PROSITE: PS01186; EGF_2; 21.
 CC PROSITE: PS00026; EGF_3; 27.
 CC PROSITE: PS01187; EGF_CA; 9.
 CC Receptor; transcription regulation; Activator; Differentiation;
 CC developmental protein; Repeat; ANK repeat; EGF-like domain;
 CC Transmembrane; Glycoprotein; Signal; Phosphorylation; Proto-oncogene.
 FT SIGNAL 1 20
 FT CHAIN 21 1964
 FT CHAIN 1411 1964
 FT CHAIN 1428 1964
 FT CHAIN 1463 1964
 FT DOMAIN 21 1443
 FT TRANSMEM 1444 1464
 FT DOMAIN 1465 1964
 FT DOMAIN 21 60
 FT DOMAIN 61 112
 FT DOMAIN 115 152
 FT DOMAIN 153 189
 FT DOMAIN 191 229
 FT DOMAIN 231 271
 FT DOMAIN 273 309
 FT DOMAIN 311 350
 FT DOMAIN 352 388
 FT DOMAIN 389 427
 FT DOMAIN 429 470
 FT DOMAIN 472 508
 FT DOMAIN 510 546
 FT DOMAIN 548 584
 FT SIGNAL 1 20
 FT CHAIN 21 1964
 FT CHAIN 1411 1964
 FT CHAIN 1428 1964
 FT CHAIN 1463 1964
 FT DOMAIN 21 1443
 FT TRANSMEM 1444 1464
 FT DOMAIN 1465 1964
 FT DOMAIN 21 60
 FT DOMAIN 61 112
 FT DOMAIN 115 152
 FT DOMAIN 153 189
 FT DOMAIN 191 229
 FT DOMAIN 231 271
 FT DOMAIN 273 309
 FT DOMAIN 311 350
 FT DOMAIN 352 388
 FT DOMAIN 389 427
 FT DOMAIN 429 470
 FT DOMAIN 472 508
 FT DOMAIN 510 546
 FT DOMAIN 548 584
 FT SIGNAL 1 20
 FT CHAIN 21 1964
 FT CHAIN 1411 1964
 FT CHAIN 1428 1964
 FT CHAIN 1463 1964
 FT DOMAIN 21 1443
 FT TRANSMEM 1444 1464
 FT DOMAIN 1465 1964
 FT DOMAIN 21 60
 FT DOMAIN 61 112
 FT DOMAIN 115 152
 FT DOMAIN 153 189
 FT DOMAIN 191 229
 FT DOMAIN 231 271
 FT DOMAIN 273 309
 FT DOMAIN 311 350
 FT DOMAIN 352 388
 FT DOMAIN 389 427
 FT DOMAIN 429 470
 FT DOMAIN 472 508
 FT DOMAIN 510 546
 FT DOMAIN 548 584
 FT SIGNAL 1 20
 FT CHAIN 21 1964
 FT CHAIN 1411 1964
 FT CHAIN 1428 1964
 FT CHAIN 1463 1964
 FT DOMAIN 21 1443
 FT TRANSMEM 1444 1464
 FT DOMAIN 1465 1964
 FT DOMAIN 21 60
 FT DOMAIN 61 112
 FT DOMAIN 115 152
 FT DOMAIN 153 189
 FT DOMAIN 191 229
 FT DOMAIN 231 271
 FT DOMAIN 273 309
 FT DOMAIN 311 350
 FT DOMAIN 352 388
 FT DOMAIN 389 427
 FT DOMAIN 429 470
 FT DOMAIN 472 508
 FT DOMAIN 510 546
 FT DOMAIN 548 584
 FT SIGNAL 1 20
 FT CHAIN 21 1964
 FT CHAIN 1411 1964
 FT CHAIN 1428 1964
 FT CHAIN 1463 1964
 FT DOMAIN 21 1443
 FT TRANSMEM 1444 1464
 FT DOMAIN 1465 1964
 FT DOMAIN 21 60
 FT DOMAIN 61 112
 FT DOMAIN 115 152
 FT DOMAIN 153 189
 FT DOMAIN 191 229
 FT DOMAIN 231 271
 FT DOMAIN 273 309
 FT DOMAIN 311 350
 FT DOMAIN 352 388
 FT DOMAIN 389 427
 FT DOMAIN 429 470
 FT DOMAIN 472 508
 FT DOMAIN 510 546
 FT DOMAIN 548 584
 FT SIGNAL 1 20
 FT CHAIN 21 1964
 FT CHAIN 1411 1964
 FT CHAIN 1428 1964
 FT CHAIN 1463 1964
 FT DOMAIN 21 1443
 FT TRANSMEM 1444 1464
 FT DOMAIN 1465 1964
 FT DOMAIN 21 60
 FT DOMAIN 61 112
 FT DOMAIN 115 152
 FT DOMAIN 153 189
 FT DOMAIN 191 229
 FT DOMAIN 231 271
 FT DOMAIN 273 309
 FT DOMAIN 311 350
 FT DOMAIN 352 388
 FT DOMAIN 389 427
 FT DOMAIN 429 470
 FT DOMAIN 472 508
 FT DOMAIN 510 546
 FT DOMAIN 548 584
 FT SIGNAL 1 20
 FT CHAIN 21 1964
 FT CHAIN 1411 1964
 FT CHAIN 1428 1964
 FT CHAIN 1463 1964
 FT DOMAIN 21 1443
 FT TRANSMEM 1444 1464
 FT DOMAIN 1465 1964
 FT DOMAIN 21 60
 FT DOMAIN 61 112
 FT DOMAIN 115 152
 FT DOMAIN 153 189
 FT DOMAIN 191 229
 FT DOMAIN 231 271
 FT DOMAIN 273 309
 FT DOMAIN 311 350
 FT DOMAIN 352 388
 FT DOMAIN 389 427
 FT DOMAIN 429 470
 FT DOMAIN 472 508
 FT DOMAIN 510 546
 FT DOMAIN 548 584
 FT SIGNAL 1 20
 FT CHAIN 21 1964
 FT CHAIN 1411 1964
 FT CHAIN 1428 1964
 FT CHAIN 1463 1964
 FT DOMAIN 21 1443
 FT TRANSMEM 1444 1464
 FT DOMAIN 1465 1964
 FT DOMAIN 21 60
 FT DOMAIN 61 112
 FT DOMAIN 115 152
 FT DOMAIN 153 189
 FT DOMAIN 191 229
 FT DOMAIN 231 271
 FT DOMAIN 273 309
 FT DOMAIN 311 350
 FT DOMAIN 352 388
 FT DOMAIN 389 427
 FT DOMAIN 429 470
 FT DOMAIN 472 508
 FT DOMAIN 510 546
 FT DOMAIN 548 584
 FT SIGNAL 1 20
 FT CHAIN 21 1964
 FT CHAIN 1411 1964
 FT CHAIN 1428 1964
 FT CHAIN 1463 1964
 FT DOMAIN 21 1443
 FT TRANSMEM 1444 1464
 FT DOMAIN 1465 1964
 FT DOMAIN 21 60
 FT DOMAIN 61 112
 FT DOMAIN 115 152
 FT DOMAIN 153 189
 FT DOMAIN 191 229
 FT DOMAIN 231 271
 FT DOMAIN 273 309
 FT DOMAIN 311 350
 FT DOMAIN 352 388
 FT DOMAIN 389 427
 FT DOMAIN 429 470
 FT DOMAIN 472 508
 FT DOMAIN 510 546
 FT DOMAIN 548 584
 FT SIGNAL 1 20
 FT CHAIN 21 1964
 FT CHAIN 1411 1964
 FT CHAIN 1428 1964
 FT CHAIN 1463 1964
 FT DOMAIN 21 1443
 FT TRANSMEM 1444 1464
 FT DOMAIN 1465 1964
 FT DOMAIN 21 60
 FT DOMAIN 61 112
 FT DOMAIN 115 152
 FT DOMAIN 153 189
 FT DOMAIN 191 229
 FT DOMAIN 231 271
 FT DOMAIN 273 309
 FT DOMAIN 311 350
 FT DOMAIN 352 388
 FT DOMAIN 389 427
 FT DOMAIN 429 470
 FT DOMAIN 472 508
 FT DOMAIN 510 546
 FT DOMAIN 548 584
 FT SIGNAL 1 20
 FT CHAIN 21 1964
 FT CHAIN 1411 1964
 FT CHAIN 1428 1964
 FT CHAIN 1463 1964
 FT DOMAIN 21 1443
 FT TRANSMEM 1444 1464
 FT DOMAIN 1465 1964
 FT DOMAIN 21 60
 FT DOMAIN 61 112
 FT DOMAIN 115 152
 FT DOMAIN 153 189
 FT DOMAIN 191 229
 FT DOMAIN 231 271
 FT DOMAIN 273 309
 FT DOMAIN 311 350
 FT DOMAIN 352 388
 FT DOMAIN 389 427
 FT DOMAIN 429 470
 FT DOMAIN 472 508
 FT DOMAIN 510 546
 FT DOMAIN 548 584
 FT SIGNAL 1 20
 FT CHAIN 21 1964
 FT CHAIN 1411 1964
 FT CHAIN 1428 1964
 FT CHAIN 1463 1964
 FT DOMAIN 21 1443
 FT TRANSMEM 1444 1464
 FT DOMAIN 1465 1964
 FT DOMAIN 21 60
 FT DOMAIN 61 112
 FT DOMAIN 115 152
 FT DOMAIN 153 189
 FT DOMAIN 191 229
 FT DOMAIN 231 271
 FT DOMAIN 273 309
 FT DOMAIN 311 350
 FT DOMAIN 352 388
 FT DOMAIN 389 427
 FT DOMAIN 429 470
 FT DOMAIN 472 508
 FT DOMAIN 510 546
 FT DOMAIN 548 584
 FT SIGNAL 1 20
 FT CHAIN 21 1964
 FT CHAIN 1411 1964
 FT CHAIN 1428 1964
 FT CHAIN 1463 1964
 FT DOMAIN 21 1443
 FT TRANSMEM 1444 1464
 FT DOMAIN 1465 1964
 FT DOMAIN 21 60
 FT DOMAIN 61 112
 FT DOMAIN 115 152
 FT DOMAIN 153 189
 FT DOMAIN 191 229
 FT DOMAIN 231 271
 FT DOMAIN 273 309
 FT DOMAIN 311 350
 FT DOMAIN 352 388
 FT DOMAIN 389 427
 FT DOMAIN 429 470
 FT DOMAIN 472 508
 FT DOMAIN 510 546
 FT DOMAIN 548 584
 FT SIGNAL 1 20
 FT CHAIN 21 1964
 FT CHAIN 1411 1964
 FT CHAIN 1428 1964
 FT CHAIN 1463 1964
 FT DOMAIN 21 1443
 FT TRANSMEM 1444 1464
 FT DOMAIN 1465 1964
 FT DOMAIN 21 60
 FT DOMAIN 61 112
 FT DOMAIN 115 152
 FT DOMAIN 153 189
 FT DOMAIN 191 229
 FT DOMAIN 231 271
 FT DOMAIN 273 309
 FT DOMAIN 311 350
 FT DOMAIN 352 388
 FT DOMAIN 389 427
 FT DOMAIN 429 470
 FT DOMAIN 472 508
 FT DOMAIN 510 546
 FT DOMAIN 548 584
 FT SIGNAL 1 20
 FT CHAIN 21 1964
 FT CHAIN 1411 1964
 FT CHAIN 1428 1964
 FT CHAIN 1463 1964
 FT DOMAIN 21 1443
 FT TRANSMEM 1444 1464
 FT DOMAIN 1465 1964
 FT DOMAIN 21 60
 FT DOMAIN 61 112
 FT DOMAIN 115 152
 FT DOMAIN 153 189
 FT DOMAIN 191 229
 FT DOMAIN 231 271
 FT DOMAIN 273 309
 FT DOMAIN 311 350
 FT DOMAIN 352 388
 FT DOMAIN 389 427
 FT DOMAIN 429 470
 FT DOMAIN 472 508
 FT DOMAIN 510 546
 FT DOMAIN 548 584
 FT SIGNAL 1 20
 FT CHAIN 21 1964
 FT CHAIN 1411 1964
 FT CHAIN 1428 1964
 FT CHAIN 1463 1964
 FT DOMAIN 21 1443
 FT TRANSMEM 1444 1464
 FT DOMAIN 1465 1964
 FT DOMAIN 21 60
 FT DOMAIN 61 112
 FT DOMAIN 115 152
 FT DOMAIN 153 189
 FT DOMAIN 191 229
 FT DOMAIN 231 271
 FT DOMAIN 273 309
 FT DOMAIN 311 350
 FT DOMAIN 352 388
 FT DOMAIN 389 427
 FT DOMAIN 429 470
 FT DOMAIN 472 508
 FT DOMAIN 510 546
 FT DOMAIN 548 584
 FT SIGNAL 1 20
 FT CHAIN 21 1964
 FT CHAIN 1411 1964
 FT CHAIN 1428 1964
 FT CHAIN 1463 1964
 FT DOMAIN 21 1443
 FT TRANSMEM 1444 1464
 FT DOMAIN 1465 1964
 FT DOMAIN 21 60
 FT DOMAIN 61 112
 FT DOMAIN 115 152
 FT DOMAIN 153 189
 FT DOMAIN 191 229
 FT DOMAIN 231 271
 FT DOMAIN 273 309
 FT DOMAIN 311 350
 FT DOMAIN 352 388
 FT DOMAIN 389 427
 FT DOMAIN 429 470
 FT DOMAIN 472 508
 FT DOMAIN 510 546
 FT DOMAIN 548 584
 FT SIGNAL 1 20
 FT CHAIN 21 1964
 FT CHAIN 1411 1964
 FT CHAIN 1428 1964
 FT CHAIN 1463 1964
 FT DOMAIN 21 1443
 FT TRANSMEM 1444 1464
 FT DOMAIN 1465 1964
 FT DOMAIN 21 60
 FT DOMAIN 61 112
 FT DOMAIN 115 152
 FT DOMAIN 153 189
 FT DOMAIN 191 229
 FT DOMAIN 231 271
 FT DOMAIN 273 309
 FT DOMAIN 311 350
 FT DOMAIN 352 388
 FT DOMAIN 389 427
 FT DOMAIN 429 470
 FT DOMAIN 472 508
 FT DOMAIN 510 546
 FT DOMAIN 548 584
 FT SIGNAL 1 20
 FT CHAIN 21 1964
 FT CHAIN 1411 1964
 FT CHAIN 1428 1964
 FT CHAIN 1463 1964
 FT DOMAIN 21 1443
 FT TRANSMEM 1444 1464
 FT DOMAIN 1465 1964
 FT DOMAIN 21 60
 FT DOMAIN 61 112
 FT DOMAIN 115 152
 FT DOMAIN 153 189
 FT DOMAIN 191 229
 FT DOMAIN 231 271
 FT DOMAIN 273 309
 FT DOMAIN 311 350
 FT DOMAIN 352 388
 FT DOMAIN 389 427
 FT DOMAIN 429 470
 FT DOMAIN 472 508
 FT DOMAIN 510 546
 FT DOMAIN 548 584
 FT SIGNAL 1 20
 FT CHAIN 21 1964
 FT CHAIN 1411 1964
 FT CHAIN 1428 1964
 FT CHAIN 1463 1964
 FT DOMAIN 21 1443
 FT TRANSMEM 1444 1464
 FT DOMAIN 1465 1964
 FT DOMAIN 21 60
 FT DOMAIN 61 112
 FT DOMAIN 115 152
 FT DOMAIN 153 189
 FT DOMAIN 191 229
 FT DOMAIN 231 271
 FT DOMAIN 273 309
 FT DOMAIN 311 350
 FT DOMAIN 352 388
 FT DOMAIN 389 427
 FT DOMAIN 429 470
 FT DOMAIN 472 508
 FT DOMAIN 510 546
 FT DOMAIN 548 584
 FT SIGNAL 1 20
 FT CHAIN 21 1964
 FT CHAIN 1411 1964
 FT CHAIN 1428 1964
 FT CHAIN 1463 1964
 FT DOMAIN 21 1443
 FT TRANSMEM 1444 1464
 FT DOMAIN 1465 1964
 FT DOMAIN 21 60
 FT DOMAIN 61 112
 FT DOMAIN 115 152
 FT DOMAIN 153 189
 FT DOMAIN 191 229
 FT DOMAIN 231 271
 FT DOMAIN 273 309
 FT DOMAIN 311 350
 FT DOMAIN 352 388
 FT DOMAIN 389 427
 FT DOMAIN 429 470
 FT DOMAIN 472 508
 FT DOMAIN 510 546
 FT DOMAIN 548 584
 FT SIGNAL 1 20
 FT CHAIN 21 1964
 FT CHAIN 1411 1964
 FT CHAIN 1428 1964
 FT CHAIN 1463 1964
 FT DOMAIN 21 1443
 FT TRANSMEM 1444 1464
 FT DOMAIN 1465 1964
 FT DOMAIN 21 60
 FT DOMAIN 61 112
 FT DOMAIN 115 152
 FT DOMAIN 153 189
 FT DOMAIN 191 229
 FT DOMAIN 231 271
 FT DOMAIN 273 309
 FT DOMAIN 311 350
 FT DOMAIN 352 388
 FT DOMAIN 389 427
 FT DOMAIN 429 470
 FT DOMAIN 472 508
 FT DOMAIN 510 546
 FT DOMAIN 548 584
 FT SIGNAL 1 20
 FT CHAIN 21 1964
 FT CHAIN 1411 1964
 FT CHAIN 1428 1964
 FT CHAIN 1463 1964
 FT DOMAIN 21 1443
 FT TRANSMEM 1444 1464
 FT DOMAIN 1465 1964
 FT DOMAIN 21 60
 FT DOMAIN 61 112
 FT DOMAIN 115 152
 FT DOMAIN 153 189
 FT DOMAIN 191 229
 FT DOMAIN 231 271
 FT DOMAIN 273 309
 FT DOMAIN 311 350
 FT DOMAIN 352 388
 FT DOMAIN 389 427
 FT DOMAIN 429 470
 FT DOMAIN 472 508
 FT DOMAIN 510 546
 FT DOMAIN 548 584
 FT SIGNAL 1 20
 FT CHAIN 21 1964
 FT CHAIN 1411 1964
 FT CHAIN 1428 1964
 FT CHAIN 1463 1964
 FT DOMAIN 21 1443
 FT TRANSMEM 1444 1464
 FT DOMAIN 1465 1964
 FT DOMAIN 21 60
 FT DOMAIN 61 112
 FT DOMAIN 115 152
 FT DOMAIN 153 189
 FT DOMAIN 191 229
 FT DOMAIN 231 271
 FT DOMAIN 273 309
 FT DOMAIN 311 350
 FT DOMAIN 352 388
 FT DOMAIN 389 427
 FT DOMAIN 429 470
 FT DOMAIN 472 508
 FT DOMAIN 510 546
 FT DOMAIN 548 584
 FT SIGNAL 1 20
 FT CHAIN 21 1964
 FT CHAIN 1411 1964
 FT CHAIN 1428 1964
 FT CHAIN 1463 1964
 FT DOMAIN 21 1443
 FT TRANSMEM 1444 1464
 FT DOMAIN 1465 1964
 FT DOMAIN 21 60
 FT DOMAIN 61 112
 FT DOMAIN 115 152
 FT DOMAIN 153 189
 FT DOMAIN 191 229
 FT DOMAIN 231 271
 FT DOMAIN 273 309
 FT DOMAIN 311 350
 FT DOMAIN 352 388
 FT DOMAIN 389 427
 FT DOMAIN 429 470
 FT DOMAIN 472 508
 FT DOMAIN 510 546
 FT DOMAIN 548 584
 FT SIGNAL 1 20
 FT CHAIN 21 1964
 FT CHAIN 1411 1964
 FT CHAIN 1428 1964
 FT CHAIN 1463 1964
 FT DOMAIN 21 1443
 FT TRANSMEM 1444 1464
 FT DOMAIN 1465 1964
 FT DOMAIN 21 60
 FT DOMAIN 61 112
 FT DOMAIN 115 152
 FT DOMAIN 153 189
 FT DOMAIN 191 229
 FT DOMAIN 231 271
 FT DOMAIN 273 309
 FT DOMAIN 311 350
 FT DOMAIN 352 388
 FT DOMAIN 389 427
 FT DOMAIN 429 470
 FT DOMAIN 472 508
 FT DOMAIN 510 546
 FT DOMAIN 548 584
 FT SIGNAL 1 20
 FT CHAIN 21 1964
 FT CHAIN 1411 1964
 FT CHAIN 1428 1964
 FT CHAIN 1463 1964
 FT DOMAIN 21 1443
 FT TRANSMEM 1444 1464
 FT DOMAIN 1465 1964
 FT DOMAIN 21 60
 FT DOMAIN 61 112
 FT DOMAIN 115 152
 FT DOMAIN 153 189
 FT DOMAIN 191 229
 FT DOMAIN 231 271
 FT DOMAIN 273 309
 FT DOMAIN 311 350
 FT DOMAIN 352 388
 FT DOMAIN 389 427
 FT DOMAIN 429 470
 FT DOMAIN 472 508
 FT DOMAIN 510 546
 FT DOMAIN 548 584
 FT SIGNAL 1 20
 FT CHAIN 21 1964
 FT CHAIN 1411 1964
 FT CHAIN 1428 1964
 FT CHAIN 1463 1964
 FT DOMAIN 21 1443
 FT TRANSMEM 1444 1464
 FT DOMAIN 1465 1964
 FT DOMAIN 21 60
 FT DOMAIN 61 112
 FT DOMAIN 115 152
 FT DOMAIN 153 189
 FT DOMAIN 191 229
 FT DOMAIN 231 271
 FT DOMAIN 273 309
 FT DOMAIN 311 350
 FT DOMAIN 352 388
 FT DOMAIN 389 427
 FT DOMAIN 429 470
 FT DOMAIN 472 508
 FT DOMAIN 510 546
 FT DOMAIN 548 584
 FT SIGNAL 1 20
 FT CHAIN 21 1964
 FT CHAIN 1411 1964
 FT CHAIN 1428 1964
 FT CHAIN 1463 1964
 FT DOMAIN 21 1443
 FT TRANSMEM 1444 1464
 FT DOMAIN 1465 1964
 FT DOMAIN 21 60
 FT DOMAIN 61 112
 FT DOMAIN 115 152
 FT DOMAIN 153 189
 FT DOMAIN 191 229
 FT DOMAIN 231 271
 FT DOMAIN 273 309
 FT DOMAIN 311 350
 FT DOMAIN 352 388
 FT DOMAIN 389 427
 FT DOMAIN 429 470
 FT DOMAIN 472 508
 FT DOMAIN 510 546
 FT DOMAIN 548 584
 FT SIGNAL 1 20
 FT CHAIN 21 1964
 FT CHAIN 1411 1964
 FT CHAIN 1428 1964
 FT CHAIN 1463 1964
 FT DOMAIN 21 1443
 FT TRANSMEM 1444 1464
 FT DOMAIN 1465 1964
 FT DOMAIN 21 60
 FT DOMAIN 61 112
 FT DOMAIN 115 152
 FT DOMAIN 153 189
 FT DOMAIN 191 229
 FT DOMAIN 231 271
 FT DOMAIN 273 309
 FT DOMAIN 311 350
 FT DOMAIN 352 388
 FT DOMAIN 389 427
 FT DOMAIN 429 470
 FT DOMAIN 472 508
 FT DOMAIN 510 546
 FT DOMAIN 548 584
 FT SIGNAL 1 20
 FT CHAIN 21 1964
 FT CHAIN 1411 1964
 FT CHAIN 1428 1964
 FT CHAIN 1463 1964
 FT DOMAIN 21 1443
 FT TRANSMEM 1444 1464
 FT DOMAIN 1

PIRSF; PIRSF002279; Notch; 1.	DR	PIRSF; PIRSF002279; Notch; 1.	DR
PRINTS; PRO0010; EGFBLLOOD.	DR	PRINTS; PRO0010; EGFBLLOOD.	DR
PRINTS; PRO0011; EGFALAMIN.	DR	PRINTS; PRO0011; EGFALAMIN.	DR
PRINTS; PRO1452; NOTCH.	DR	PRINTS; PRO1452; NOTCH.	DR
SMART; SM00248; ANK; 6.	DR	SMART; SM00248; ANK; 6.	DR
SMART; SM00179; EGF_CA; 25.	DR	SMART; SM00179; EGF_CA; 25.	DR
SMART; SM00004; NL; 2.	DR	SMART; SM00004; NL; 2.	DR
PROSITE; PS50297; ANK_REP_REGION; 1.	DR	PROSITE; PS50297; ANK_REP_REGION; 1.	DR
PROSITE; PS50088; ANK_REPEAT; 4.	DR	PROSITE; PS50088; ANK_REPEAT; 4.	DR
PROSITE; PS00010; ASX_HYDROXYL; 22.	DR	PROSITE; PS00010; ASX_HYDROXYL; 22.	DR
PROSITE; PS00022; EGF_1; 35.	DR	PROSITE; PS00022; EGF_1; 35.	DR
PROSITE; PS01186; EGF_2; 26.	DR	PROSITE; PS01186; EGF_2; 26.	DR
PROSITE; PS50026; EGF_3; 36.	DR	PROSITE; PS50026; EGF_3; 36.	DR
PROSITE; PS0187; EGF_CA; 21.	DR	PROSITE; PS0187; EGF_CA; 21.	DR
Receptor; Transcription regulation; Activator; Differentiation; Developmental protein; Repeat; ANK repeat; EGF-like domain; Transmembrane; Glycoprotein; Signal; Phosphorylation.	KW	Receptor; Transcription regulation; Activator; Differentiation; Developmental protein; Repeat; ANK repeat; EGF-like domain; Transmembrane; Glycoprotein; Signal; Phosphorylation.	KW
SIGNAL	1	SIGNAL	18
FT CHAIN	19	FT CHAIN	2531
FT CHAIN	1711	FT CHAIN	2531
FT CHAIN	1744	FT CHAIN	2531
DOMAIN	19	DOMAIN	1723
TRANSMEM	1724	TRANSMEM	1746
DOMAIN	1747	DOMAIN	2531
DOMAIN	20	DOMAIN	58
DOMAIN	59	DOMAIN	99
DOMAIN	102	DOMAIN	139
DOMAIN	140	DOMAIN	176
DOMAIN	178	DOMAIN	216
DOMAIN	218	DOMAIN	255
DOMAIN	257	DOMAIN	293
DOMAIN	295	DOMAIN	333
DOMAIN	335	DOMAIN	371
DOMAIN	372	DOMAIN	410
DOMAIN	412	DOMAIN	450
DOMAIN	452	DOMAIN	488
DOMAIN	490	DOMAIN	526
DOMAIN	528	DOMAIN	564
DOMAIN	566	DOMAIN	601
DOMAIN	603	DOMAIN	639
DOMAIN	641	DOMAIN	676
DOMAIN	678	DOMAIN	714
DOMAIN	716	DOMAIN	751
DOMAIN	753	DOMAIN	789
DOMAIN	791	DOMAIN	827
DOMAIN	829	DOMAIN	867
DOMAIN	869	DOMAIN	905
DOMAIN	907	DOMAIN	943
DOMAIN	945	DOMAIN	981
DOMAIN	983	DOMAIN	1019
DOMAIN	1021	DOMAIN	1057
DOMAIN	1059	DOMAIN	1095
DOMAIN	1097	DOMAIN	1143
DOMAIN	1145	DOMAIN	1181
DOMAIN	1183	DOMAIN	1219
DOMAIN	1221	DOMAIN	1265
DOMAIN	1267	DOMAIN	1305
DOMAIN	1307	DOMAIN	1346
DOMAIN	1348	DOMAIN	1384
DOMAIN	1387	DOMAIN	1426
REPEAT	1445	REPEAT	1480
REPEAT	1481	REPEAT	1522
REPEAT	1523	REPEAT	1562
REPEAT	1597	REPEAT	1946
REPEAT	1950	REPEAT	1980
REPEAT	1984	REPEAT	2013
REPEAT	2017	REPEAT	2046
REPEAT	2050	REPEAT	2079
DOMAIN	1730	DOMAIN	1733
DOMAIN	1891	DOMAIN	1894
DOMAIN	2258	DOMAIN	2261

FT DOMAIN 2497 2500 POLY-SER. CC
 FT SITE 1654 1655 CLEAVAGE BY (FURIN-LIKE PROTEASE) (BY CC
 FT BY SIMILARITY. CC
 FT DISULFID 24 37 BY SIMILARITY. CC
 FT DISULFID 31 46 BY SIMILARITY. CC
 FT DISULFID 48 57 BY SIMILARITY. CC
 FT DISULFID 63 74 BY SIMILARITY. CC
 FT DISULFID 68 87 BY SIMILARITY. CC
 FT DISULFID 89 98 BY SIMILARITY. CC
 FT DISULFID 106 117 BY SIMILARITY. CC
 FT DISULFID 111 127 BY SIMILARITY. CC
 FT DISULFID 129 138 BY SIMILARITY. CC
 FT DISULFID 144 155 BY SIMILARITY. CC
 FT DISULFID 149 164 BY SIMILARITY. CC
 FT DISULFID 166 175 BY SIMILARITY. CC
 FT DISULFID 182 195 BY SIMILARITY. CC
 FT DISULFID 189 204 BY SIMILARITY. CC
 FT DISULFID 206 215 BY SIMILARITY. CC
 FT DISULFID 222 233 BY SIMILARITY. CC
 FT DISULFID 227 243 BY SIMILARITY. CC

Query Match 12.7%; Score 95.5; DB 1; Length 2531;
 Best Local Similarity 26.8%; Pred. No. 0.2;
 Matches 34; Conservative 7; Mismatches 35; Indels 51; Gaps 7;

QY 15 ISACMLLCHGSLQHTFQHHLRPEGTCE-VIAAHC-----CNKN 56
 DB 679 IDECAGSPCH-----NGGTCEGDIAGFTCRPGYHDPCTCLSEVNECSN 723

QY 57 -----RIERSQTVKCSCLPGKVGATRNRPSCVDASIVIGKWCMEPCLEGECKTLP 111
 DB 724 PCTHGACRDLGKYGKDCAFG-WSGTNCD-----INNCEENPCVNGTCKDM- 771

QY 112 DNSGWMVC 118
 DB 772 -TSGYVC 777

RESULT 5
 CRB DROME STANDARD; PRT; 2139 AA.
 AC P10040;
 DT 01-MAR-1989 (Rel. 10, Created)
 DT 01-MAY-1991 (Rel. 18, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Crumbs protein precursor (95F).
 GN CRB.
 OS Drosophila melanogaster (Fruit fly).
 OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
 OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
 OC Ephydroidea; Drosophilidae; Drosophila.
 OX NCBI_TaxID=7227;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Oregon-R; TISSUE=Embryo;
 RX MEDLINE=90263104; PubMed=2344615;
 RA Tepass U., Theres C., Knust E.;
 RT "Crumbs encodes an EGF-like protein expressed on apical membranes of
 RT Drosophila epithelial cells and required for organization of
 RT epithelia.";
 RL Cell 61:787-799(1990).
 RN [2]
 RP SEQUENCE OF 1663-1955 FROM N.A.
 RC TISSUE=Embryo;
 RX MEDLINE=87218537; PubMed=3107986;
 RA Knust E., Dietrich U., Tepass U., Bremer K.A., Weigel D.,
 RA Vaessin H., Campos-Ortega J.A.;
 RT "EGF homologous sequences encoded in the genome of Drosophila
 RT melanogaster, and their relation to neurogenic genes.";
 RL EMBO J. 6:761-766(1987).
 CC -!- FUNCTION: May play a role in the development of epithelia,
 CC possibly for the establishment and/or maintenance of cell
 CC polarity. It may act as a signal.

CC -!- SUBCELLULAR LOCATION: Type I membrane protein.
 CC -!- PTM: PHOSPHORYLATED IN THE CYTOPLASMIC DOMAIN (POTENTIAL).
 CC -!- SIMILARITY: Contains 29 EGF-like domains.
 CC -!- SIMILARITY: Contains 3 laminin G-like domains.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; M33753; AAA28428.1; ALT_SEQ.
 DR EMBL; X05144; CAA28793.1; -.
 DR PIR; A35672; A35672.
 DR PIR; B26637; B26637.
 DR HSSP; P00740; 1EDM.
 DR FlyBase; FBgn000368; crb.
 DR GO; GO:0016324; C:apical plasma membrane; NAS.
 DR GO; GO:0016327; C:apicolateral plasma membrane; IDA.
 DR GO; GO:0007163; P:establishment and/or maintenance of cell po...; IMP.
 DR GO; GO:0016332; P:establishment and/or maintenance of polarit...; IMP.
 DR GO; GO:0016334; P:establishment and/or maintenance of polarit...; IMP.
 DR GO; GO:0045494; P:photoreceptor maintenance; IMP.
 DR GO; GO:0042052; P:rhabdome development; NAS.
 DR GO; GO:0045186; P:zonula adherens assembly; IMP.
 DR GO; GO:0045218; P:zonula adherens maintenance; IMP.
 DR InterPro; IPR000152; Asx hydroxyl S.
 DR InterPro; IPR008985; ConA like lec_gl.
 DR InterPro; IPR000742; EGF_2.
 DR InterPro; IPR001881; EGF_Ca.
 DR InterPro; IPR001438; EGF_II.
 DR InterPro; IPR006209; EGF like.
 DR InterPro; IPR002049; laminin EGF.
 DR InterPro; IPR001791; Laminin_G.
 DR Pfam; PF00008; EGF; 26.
 DR Pfam; PF00054; laminin G; 3.
 DR PRINTS; PR00010; EGFBL000.
 DR PRINTS; PR00011; EGF_LAMININ.
 DR SMART; SM00179; EGF_CA; 11.
 DR SMART; SM00282; LamG; 3.
 DR PROSITE; PS00010; ASX HYDROXYL; 15.
 DR PROSITE; PS00022; EGF_1; 26.
 DR PROSITE; PS01186; EGF_2; 17.
 DR PROSITE; PS00026; EGF_3; 27.
 DR PROSITE; PS01187; EGF_CA; 12.
 DR PROSITE; PS00025; LAM G DOMAIN; 3.
 DR Differentiation; Repeat; EGF-like domain; Transmembrane;
 KW Glycoprotein; Signal; Phosphorylation.
 FT SIGNAL 1 90
 FT CHAIN 91 2139 CRUMBS PROTEIN.
 FT DOMAIN 91 2084 EXTRACELLULAR (POTENTIAL).
 FT TRANSMEM 2085 2111 POTENTIAL.
 FT DOMAIN 2112 2139 CYTOPLASMIC (POTENTIAL).
 FT DOMAIN 267 303 EGF-LIKE 1.
 FT DOMAIN 306 343 EGF-LIKE 2.
 FT DOMAIN 348 386 EGF-LIKE 3.
 FT DOMAIN 388 425 EGF-LIKE 4.
 FT DOMAIN 427 463 EGF-LIKE 5.
 FT DOMAIN 464 500 EGF-LIKE 6.
 FT DOMAIN 501 532 EGF-LIKE 7.
 FT DOMAIN 545 581 EGF-LIKE 8.
 FT DOMAIN 582 611 EGF-LIKE 9.
 FT DOMAIN 609 646 EGF-LIKE 10.
 FT DOMAIN 648 685 EGF-LIKE 11.
 FT DOMAIN 687 723 EGF-LIKE 12.
 FT DOMAIN 725 761 EGF-LIKE 13.
 FT DOMAIN 763 800 EGF-LIKE 14.
 FT DOMAIN 802 838 EGF-LIKE 15.
 FT DOMAIN 840 902 EGF-LIKE 16.
 FT DOMAIN 904 940 EGF-LIKE 17.
 FT DOMAIN 942 978 EGF-LIKE 18, CALCIUM-BINDING (POTENTIAL).

InterPro; IPR000152; Asx hydroxyl_s.
 L InterPro; IPR001774; DSL.
 DR InterPro; IPR000742; EGF 2.
 DR InterPro; IPR001891; EGF_Ca.
 DR InterPro; IPR001438; EGF_II.
 DR InterPro; IPR006209; EGF_Like.
 DR InterPro; IPR002049; Laminin EGF.
 DR InterPro; IPR009041; PMP inhibitor.
 DR InterPro; IPR001007; VWF_C.
 DR Pfam; PF01414; DSL; 1.
 DR Pfam; PF00008; EGF; 14.
 DR PRINTS; PRO0010; EGFBLD.
 DR PRINTS; PRO0011; EGF_LAMININ.
 DR SMART; SM00051; DSL; 1.
 DR SMART; SM00179; EGF_CA; 10.
 DR SMART; SM00214; VWC; 1.
 DR PROSITE; PS00010; ASX HYDROXYL; 10.
 DR PROSITE; PS00022; EGF_1; 16.
 DR PROSITE; PS01186; EGF_2; 12.
 DR PROSITE; PS00026; EGF_3; 15.
 DR PROSITE; PS01187; EGF_CA; 8.
 DR PROSITE; PS01208; VWF_C 1; FALSE NEG.
 KW Calcium-binding; EGF-like domain; Developmental protein;
 KW Calcium-binding; EGF-like domain; Glycoprotein; Signal.
 KW Repeat; Transmembrane; Signal.
 FT SIGNAL 1 26 POTENTIAL.
 FT CHAIN 27 1213 JAGGED 3.
 FT DOMAIN 27 1064 EXTRACELLULAR (POTENTIAL).
 FT TRANSMEM 1065 1087 POTENTIAL.
 FT DOMAIN 1088 1213 CYTOPLASMIC (POTENTIAL).
 FT DOMAIN 164 226 DSL.
 FT DOMAIN 227 260 EGF-LIKE 1.
 FT DOMAIN 258 291 EGF-LIKE 2.
 FT DOMAIN 293 331 EGF-LIKE 3.
 FT DOMAIN 333 369 EGF-LIKE 4.
 FT DOMAIN 371 407 EGF-LIKE 5.
 FT DOMAIN 409 445 EGF-LIKE 6.
 FT DOMAIN 447 482 EGF-LIKE 7.
 FT DOMAIN 484 520 EGF-LIKE 8.
 FT DOMAIN 522 558 EGF-LIKE 9.
 FT DOMAIN 592 624 EGF-LIKE 10.
 FT DOMAIN 626 662 EGF-LIKE 11.
 FT DOMAIN 664 700 EGF-LIKE 12.
 FT DOMAIN 702 738 EGF-LIKE 13.
 FT DOMAIN 746 777 EGF-LIKE 14.
 FT DOMAIN 779 815 EGF-LIKE 15.
 FT DOMAIN 817 853 EGF-LIKE 16.
 FT DOMAIN 860 914 VWF_C.
 FT DOMAIN 918 956 EGF-LIKE 17.
 FT DISULFID 231 242 BY SIMILARITY.
 FT DISULFID 235 248 BY SIMILARITY.
 FT DISULFID 250 259 BY SIMILARITY.
 FT DISULFID 262 273 BY SIMILARITY.
 FT DISULFID 268 279 BY SIMILARITY.
 FT DISULFID 281 290 BY SIMILARITY.
 FT DISULFID 297 309 BY SIMILARITY.
 FT DISULFID 303 319 BY SIMILARITY.
 FT DISULFID 321 330 BY SIMILARITY.
 FT DISULFID 337 348 BY SIMILARITY.
 FT DISULFID 342 357 BY SIMILARITY.
 FT DISULFID 359 368 BY SIMILARITY.
 FT DISULFID 375 386 BY SIMILARITY.
 FT DISULFID 380 395 BY SIMILARITY.
 FT DISULFID 397 406 BY SIMILARITY.
 FT DISULFID 413 424 BY SIMILARITY.
 FT DISULFID 418 433 BY SIMILARITY.
 FT DISULFID 435 444 BY SIMILARITY.
 FT DISULFID 451 461 BY SIMILARITY.
 FT DISULFID 455 470 BY SIMILARITY.
 FT DISULFID 472 481 BY SIMILARITY.
 FT DISULFID 488 499 BY SIMILARITY.
 FT DISULFID 493 508 BY SIMILARITY.
 FT DISULFID 510 519 BY SIMILARITY.
 FT DISULFID 526 537 BY SIMILARITY.

FT DISULFID 531 546 BY SIMILARITY.
 FT DISULFID 548 557 BY SIMILARITY.
 FT DISULFID 630 641 BY SIMILARITY.
 FT DISULFID 635 650 BY SIMILARITY.
 FT DISULFID 652 661 BY SIMILARITY.
 FT DISULFID 668 679 BY SIMILARITY.
 FT DISULFID 673 688 BY SIMILARITY.
 FT DISULFID 690 699 BY SIMILARITY.
 FT DISULFID 706 717 BY SIMILARITY.
 FT DISULFID 711 726 BY SIMILARITY.
 FT DISULFID 728 737 BY SIMILARITY.
 FT DISULFID 783 794 BY SIMILARITY.
 FT DISULFID 788 803 BY SIMILARITY.
 FT DISULFID 805 814 BY SIMILARITY.
 FT DISULFID 821 832 BY SIMILARITY.
 FT DISULFID 826 841 BY SIMILARITY.
 FT DISULFID 843 852 BY SIMILARITY.
 FT DOMAIN 938 941 POLY-PRO.
 FT CARBOHYD 119 139 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 214 234 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 556 566 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 742 742 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 957 957 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 988 988 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 1042 1042 N-LINKED (GLCNAC. .) (POTENTIAL).
 SQ SEQUENCE 1213 AA; 133365 MW; 5C5F16A7E20D9534 CRC64;
 Query Match 12.1%; Score 91.5; DB 1; Length 1213;
 Best Local Similarity 24.8%; Pred. No. 0.23;
 Matches 35; Conservative 19; Mismatches 50; Indels 37; Gaps 8;
 QY 17 ACAMLLC--HGSLOHTFQOHLHRP---EGGTCEV---IAAHRCCNKNRIBERSQTVKC 67
 DB 336 ACLSNPCANGCTCKTSQGYECHCAIGWSGTSCINVDCTPNQCKHGTCQDLVNGFKC 395
 QY 68 SCLP---GKV---AGTTNRNRCSCVDASI---VIGKWCMEP-----100
 DB 396 ACPPHWGTGTCQIDANECEDEXP-CVNAKSCHNLIGAYFCECLPFGWSGQNCINNDCKGQ 454
 QY 101 CLEGECKTLPDNGWMCATG 121
 DB 455 CLNGGTCKDLVNGYRCLCPPG 475
 RESULT 7
 DP87_DP87 ID DP87_DP87 STANDARD; PRT; 555 AA.
 AC Q04503;
 DT 01-FEB-1994 (Rel. 28, Created)
 DT 01-FEB-1994 (Rel. 28, Last sequence update)
 DT 15-MAR-2004 (Rel. 43, Last annotation update)
 DE Prespore protein DP87 precursor.
 GN COTD OR DP87.
 OS Dictyostelium discoideum (Slime mold).
 OC Eukaryota; Mycetozoa; Dictyosteliida; Dictyostelium.
 OX NCBI_TaxID=44689;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=AX3;
 RX MEDLINE=94008533; PubMed=8404532;
 RA Ozaki T., Nakao H., Orii H., Morio T., Takeuchi I., Tasaka M.;
 RT "Developmental regulation of transcription of a novel
 RL Prespore-specific gene (Op87) in Dictyostelium discoideum";
 CC Development 117:1299-1308(1993).
 CC -[- SUBCELLULAR LOCATION: Stored in prespore vacuoles until it is
 CC discharged into the interspace of spores during spore formation.
 CC -[- INDUCTION: By exogenous cAMP, repressed by DIF.
 CC -[- SIMILARITY: Contains 7 prespore motif repeats.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way

CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

DR EMBL; D13973; BAA03083.1; --
 DR DictyBase; DDB0001389; --
 DR InterPro; IPR007643; Dict_spore_N.
 DR InterPro; IPR003645; FOLN.
 DR Pfam; PF04562; Dict_spore_N; 1.
 DR SMART; SM00274; FOLN; 6.
 KW Glycoprotein; Sporulation; Signal; Repeat.
 FT SIGNAL 1 17 POTENTIAL.
 FT CHAIN 18 555 PRESPORE PROTEIN DP87.
 FT REPEAT 149 161 PRESPORE MOTIF 1.
 FT REPEAT 176 188 PRESPORE MOTIF 2.
 FT REPEAT 210 222 PRESPORE MOTIF 3.
 FT REPEAT 241 253 PRESPORE MOTIF 4.
 FT REPEAT 305 317 PRESPORE MOTIF 5.
 FT REPEAT 333 345 PRESPORE MOTIF 6.
 FT REPEAT 363 375 PRESPORE MOTIF 7.
 FT DOMAIN 438 473 THR-RICH.
 FT DOMAIN 476 555 ALA/SER-RICH.
 FT CARBOHYD 103 103 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 306 306 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 418 418 N-LINKED (GLCNAC. . .) (POTENTIAL).
 SQ SEQUENCE 555 AA; 58673 MW; 6DA8A6D1C36F8DD CRC64;

Query Match 12.1%; Score 91; DB 1; Length 555;
 Best Local Similarity 22.9%; Pred. No. 0.12;
 Matches 27; Conservative 14; Mismatches 53; Indels 24; Gaps 5;

QY 34 HHLHR-----PEGGTCEVIAHRCCKNRIBERSQTV-----KCSCLPGKVAGTTRNRPSC 84
 Db 238 HNLCRFGCEGSHCEVLEKHPVCRNHNPPPPPPQICGNSVCGGYCTTINHGHTC 297
 QY 85 VDASIVIGKWCMEPCLEGECKTLTP-----DNSGW-MCATGNKIKTRIHP 131
 Db 298 IRGD----GYLCNQTRCPHDYQCTETISNVKSPKNDCKWRCPPGSSCFNSRNGP 351

RESULT 8

NOTC DROME STANDARD; PRT; 2703 AA.
 AC P07207; O97458; P04154; O9W478;
 DT 01-NOV-1986 (Rel. 03, Created)
 DT 28-FEB-2003 (Rel. 41, Last sequence update)
 DT 15-MAR-2004 (Rel. 43, Last annotation update)
 DE Neurogenic locus Notch protein precursor.
 GN N OR EG:140G11.1 OR EG:163A10.2 OR CG3936.
 OS Drosophila melanogaster (Fruit fly).
 OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
 OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
 OC Ephydroidea; Drosophilidae; Drosophila.
 OX NCBI_TaxID=7227;
 [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Oregon-R; TISSUE=Embryo;
 RX MEDLINE=86079539; PubMed=3935325;
 RA Wharton K.A., Johansen K.M., Xu T., Artavanis-Tsakonas S.;
 RT "Nucleotide sequence from the neurogenic locus notch implies a gene
 RT product that shares homology with proteins containing EGF-like
 RT repeats.";
 RL Cell 43:567-581(1985).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Canton-S, and Oregon-R; TISSUE=Embryo;
 RX MEDLINE=87064624; PubMed=3097517;
 RA Kidd S., Kelley M.R., Young M.W.;
 RT "Sequence of the notch locus of Drosophila melanogaster: relationship
 RT of the encoded protein to mammalian clotting and growth factors.";
 RL Mol. Cell. Biol. 6:3094-3108(1986).
 RN [3]
 RP SEQUENCE FROM N.A.

RC STRAIN=Berkeley;
 RX MEDLINE=20196006; PubMed=10731132;
 RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
 RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Gallie R.F.,
 RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
 RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
 RA Brandon R.C., Rogers Y.-H.C., Blake J.R.G., Champ M., Pfeiffer B.D.,
 RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.D.G.,
 RA Abril J.F., Agbayani A., An H.-J., Andrews-Fannkuch C., Baldwin D.,
 RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
 RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
 RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brotter P.,
 RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
 RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
 RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
 RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
 RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
 RA Fosler C., Gabriellian A.E., Garg N.S., Gelbart W.M., Glasser K.,
 RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
 RA Harris N.L., Harvey D.A., Heiman T.J., Hernandez J.R., Houck J.,
 RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
 RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
 RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
 RA Lasko P., Lei Y., Levitsky A.A., Li J.H., Li Z., Liang Y., Lin X.,
 RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
 RA Markulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
 RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
 RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacleb J.M.,
 RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
 RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
 RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
 RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
 RA Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
 RA Wang Z.-Y., Wasserman D.A., Weinstein G.M., Weissbach J.,
 RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
 RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
 RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
 RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
 RT "The genome sequence of Drosophila melanogaster.";
 RL Science 287:2185-2195 (2000).
 RN [4]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Oregon-R;
 RX MEDLINE=20196011; PubMed=10731137;
 RA Benos P.V., Gatt M.K., Ashburner M., Murphy L., Harris D.,
 RA Barrell B.G., Ferraz C., Vidal S., Brun C., Demallies J., Cadieu E.,
 RA Dreano S., Gloux S., Lelaure V., Mottier S., Galibert F., Borkova D.,
 RA Minana B., Kafatos F.C., Louis C., Siden-Kiamos I., Bolshakov S.,
 RA Papagiannakis G., Spanos L., Cox S., Madueno E., de Pablos B.,
 RA Mordell J., Peter A., Schoettler P., Werner M., Mourikioti F.,
 RA Beinart N., Dowe G., Schaefer U., Jaekle H., Bucheton A.,
 RA Callister D.M., Campbell L.A., Darlamitson A., Henderson N.S.,
 RA McMillan P.J., Salles C., Tait E.A., Valenti P., Saunders R.D.C.,
 RA Glover D.M.;
 RT "From sequence to chromosome: the tip of the X chromosome of D.
 RT melanogaster.";
 RL Science 287:2220-2222 (2000).
 RN [5]
 RP SEQUENCE OF 2505-2611 FROM N.A.
 RX MEDLINE=85099329; PubMed=2981631;
 RA Wharton K.A., Vedvobnick B., Finnerty V.G., Artavanis-Tsakonas S.;
 RT "opa: a novel family of transcribed repeats shared by the Notch locus
 RT and other developmentally regulated loci in D. melanogaster.";
 RL Cell 40:55-62(1985).
 RN [6]
 RP SEQUENCE OF 1-8 FROM N.A.
 RX MEDLINE=87257846; PubMed=3037327;
 RA Kelley M.R., Kidd S., Berg R.L., Young M.W.;
 RT "Restriction of P-element insertions at the Notch locus of Drosophila
 RT melanogaster.";
 RL Mol. Cell. Biol. 7:1545-1548(1987).
 RN [7]
 RP INTERACTION WITH DX, AND MUTANT SU42C.

RX MEDLINE=94215489; PubMed=8162848;
RA Diederich R.J., Matsumo K., Hing H., Artavanis-Tsakonas S.;
RT "Cytosolic interaction between deltex and Notch ankyrin repeats
RL implicates deltex in the Notch signaling pathway.";
RN Development 120:473-481(1994).
[8]
RN INTERACTION WITH DX.
RX MEDLINE=95401878; PubMed=7671825;
RA Matsumo K., Diederich R.J., Go M.J., Blaumueller C.M.,
RA Artavanis-Tsakonas S.;
RT "Deltex acts as a positive regulator of Notch signaling through
RL interactions with the Notch ankyrin repeats.";
RN Development 121:2633-2644(1995).
[9]
RN S3 CLEAVAGE BY PSN.
RX MEDLINE=99221487; PubMed=10206646;
RA Struhl G., Greenwald I.;
RT "Presenilin is required for activity and nuclear access of Notch in
RL Drosophila.";
RN Nature 398:522-525(1999).
[10]
RN S3 CLEAVAGE BY PSN.
RX MEDLINE=99221488; PubMed=10206647;
RA Ye Y., Lukinova N., Fortini M.E.;
RT "Neurogenic phenotypes and altered Notch processing in Drosophila
RL Presenilin mutants.";
RN Nature 398:525-529(1999).
[11]
RN S2 CLEAVAGE BY KUZ.
RX MEDLINE=21657146; PubMed=11799064;
RA Lieber T., Kidd S., Young M.W.;
RT "kuzbanian-mediated cleavage of Drosophila Notch.";
RL Genes Dev. 16:209-221(2002).
[12]
RN MUTANT MCDS.
RX MEDLINE=21575956; PubMed=11719214;
RA Romain P., Khechumian K., Seugnet L., Arbogast N., Ackermann C.,
RA Heitzler P.;
RT "Novel Notch alleles reveal a Deltex-dependent pathway repressing
RL neural fate.";
RN Curr. Biol. 11:1729-1738(2001).
[13]
RN REVIEW.
RX MEDLINE=22256570; PubMed=12369105;
RA Fortin P.;
RT "General outlines of the molecular genetics of the Notch signalling
RL pathway in Drosophila melanogaster: a review.";
RN Hereditas 136:89-96(2002).
CC -!- FUNCTION: Signaling protein, which regulates, with both positive
CC and negative signals, the differentiation of at least central and
CC peripheral nervous system and eye, wing disk, oogenesis, segmental
CC appendages such as antennae and legs, and muscles, through lateral
CC inhibition or induction. Functions as a receptor for membrane-
CC bound ligands Delta and Serrate to regulate cell-fate
CC determination. Upon ligand activation and releasing from the cell
CC membrane, the Notch intracellular domain (NICD) forms a
CC transcriptional activator complex with Su(H) (Suppressor of
CC hairless) and activates genes of the E(spl) complex. Essential for
CC proper differentiation of ectoderm.
CC -!- SUBUNIT: Interacts with Su(H) when activated. Interacts with Dx
CC via its ANK repeats.
CC -!- SUBCELLULAR LOCATION: Type I membrane protein. Upon activation and
CC S3 cleavage, it is released from the cell membrane and enters into
CC the nucleus in conjunction with Su(H).
CC -!- PTM: Upon binding its ligands such as Delta or Serrate, it is
CC cleaved (S2 cleavage) in its extracellular domain, close to the
CC transmembrane domain. S2 cleavage is probably mediated by Kuz. It
CC is then cleaved (S3 cleavage) downstream of its transmembrane
CC domain, releasing it from the cell membrane. S3 cleavage requires
CC Psn.
CC -!- SIMILARITY: Belongs to the NOTCH family.
CC -!- SIMILARITY: Contains 36 EGF-like domains.
CC -!- SIMILARITY: Contains 3 Lin/Notch repeats.

CC -!- SIMILARITY: Contains 6 ANK repeats.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; M16152; AAB59220.1; -;
DR EMBL; M16153; AAB59220.1; JOINED.
DR EMBL; M16149; AAB59220.1; JOINED.
DR EMBL; M16150; AAB59220.1; JOINED.
DR EMBL; M16151; AAB59220.1; JOINED.
DR EMBL; K03508; AAA28725.1; -;
DR EMBL; M13689; AAA28725.1; JOINED.
DR EMBL; K03507; AAA28725.1; JOINED.
DR EMBL; A003426; AAF45848.2; -;
DR EMBL; AL035436; CAB37610.1; -;
DR EMBL; AL035395; CAB37610.1; JOINED.
DR EMBL; M12175; AAA74496.1; -;
DR EMBL; M16025; AAA28726.1; -;
CC -----
CC Query Match 11.5%; Score 86.5; DB 1; Length 2703;
CC Best Local Similarity 25.9%; Pred. No. 1.7;
CC Matches 22; Conservative 10; Mismatches 34; Indels 19; Gaps 3;
CC -----
CC QY 40 EGGTCV-----IAAHRCCNKNRIERSQTVKCSCLPGKVA-----GTTNRPS 83
CC 1214 QQCNELNIDDCAPNCPQNGGTCRDMVNFSCFPGTGTGIIICLNKDKCKPGACINNGS 1273
CC QY 84 CVDASIVIGKWCMEPEPCLEGBECK 108
CC 1274 CIDR---VGGFECVCQPGFVGARCE 1295
CC -----
CC RESULT 9
CC DLL4_HUMAN
CC ID DLL4_HUMAN STANDARD; PRT; 685 AA.
CC AC Q9NR61; Q9NOT9;
CC DT 16-OCT-2001 (Rel. 40, Created)
CC DT 16-OCT-2001 (Rel. 40, Last sequence update)
CC DT 15-MAR-2004 (Rel. 43, Last annotation update)
CC DE Delta-like protein 4 precursor (Drosophila Delta homolog 4)
CC DE (UNQ1895/PRO4341).
CC GN DLL4.
CC OS Homo sapiens (Human).
CC OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
CC OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
CC RN [1] NCBI_TaxID=9606;
CC RP SEQUENCE FROM N.A.
CC RC TISSUE=Brain;
CC RX MEDLINE=20296719; PubMed=10837024;
CC RA Shutter J.R., Scully S., Pan W., Richards W.G., Kitajewski J.,
CC Deblandre G.A., Kintner C.R., Stark K.B.;
CC RT "Dl4, a novel Notch ligand expressed in arterial endothelium.";
CC RL Genes Dev. 14:1313-1318(2000).
CC RN [2]
CC RP SEQUENCE FROM N.A.
CC RA Sakano S.;
CC RL Submitted (JAN-2000) to the EMBL/GenBank/DBJ databases.
CC RN [3]
CC RP SEQUENCE FROM N.A.
CC RA Yoneya T., Tahara T., Nagao K., Yamada Y., Yamamoto T., Miyatani S.,
CC Nishikawa M.;
CC RT "Molecular cloning of Delta-4, a new mouse and human Notch ligand.";
CC RL Submitted (MAY-2000) to the EMBL/GenBank/DBJ databases.
CC RN [4]
CC RP SEQUENCE FROM N.A.
CC RX MEDLINE=22887296; PubMed=12975309;
CC Clark H.F., Gurney A.L., Abaya E., Baker K., Baldwin D., Brush J.,

RA Chen J., Chow B., Chui C., Crowley C., Currell B., Deuel B., Dowd P.,
 RA Eaton D., Foster J., Grimaldi C., Gu Q., Hass P.E., Heidens S.,
 RA Huang A., Kim H.S., Klimowski L., Jin Y., Johnson S., Lee J.,
 RA Lewis L., Liao D., Mark M., Robbie E., Sanchez C., Schoenfeld J.,
 RA Seshagiri S., Simmons L., Singh J., Smith V., Stinson J., Vagts A.,
 RA Vandlen R., Watanabe C., Wiand D., Woods K., Xie M.-H., Yansura D.,
 RA Yi S., Yu G., Yuan J., Zhang M., Zhang Z., Goddard A., Wood W.I.,
 RA Godowski P.;
 RT "The secreted protein discovery initiative (SPDI), a large-scale
 RT effort to identify novel human secreted and transmembrane proteins: a
 RT bioinformatics assessment.";
 RL Genome Res. 13:2265-2270(2003).
 RN [5]
 RP SEQUENCE OF 33-685 FROM N.A.
 RC
 RA TISSUE=Placenta;
 RA Malmos C., Modlich U., Lewis J., Harris A., Bicknell R.,
 RA Ish-Horowicz D.;
 RT "A novel Delta gene expressed in embryonic and tumour vasculature.";
 RT Submitted (JUN-2000) to the EMBL/GenBank/DBJ databases.
 RC !- FUNCTION: Plays a role in the Notch signaling pathway. Activates
 CC Notch-1 and Notch-4 (By similarity).
 CC !- SUBUNIT: Binds to Notch-1 and Notch-4 (By similarity).
 CC !- SUBCELLULAR LOCATION: Type I membrane protein (Probable).
 CC !- TISSUE SPECIFICITY: Expressed in vascular endothelium.
 CC !- DOMAIN: The Delta-Serrate-Lag2 (DSL) domain is required for
 CC binding to the Notch receptor.
 CC !- SIMILARITY: Contains 8 EGF-like domains.
 CC !- SIMILARITY: Contains 1 DSL domain.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 CC EMBL: AF253468; AAF76427.1; -
 CC EMBL: AB036931; BAB16085.1; -
 CC EMBL: AB043894; BAB18581.1; -
 CC EMBL: AY358894; AAQ89253.1; -
 CC EMBL: AF279305; AAF81912.1; -
 CC F01; JC7570; JC7570.
 CC HSP; P00740; IEDM.
 CC Genew; HGNC:2910; DLL4.
 CC MIM: 605185; -
 CC GO: GO:0005112; F:Notch binding; TAS.
 CC GO: GO:0006015; P:circulation; TAS.
 CC GO: GO:0007165; P:signal transduction; TAS.
 CC InterPro; IPR000152; Asx_hydroxyl_S.
 CC InterPro; IPR001774; DSL.
 CC InterPro; IPR000742; EGF_2.
 CC InterPro; IPR001881; EGF_Ca.
 CC InterPro; IPR001438; EGF_I1.
 CC InterPro; IPR006209; EGF_Like.
 CC InterPro; IPR002049; Laminin_EGF.
 CC Pfam; PF01414; DSL; 1.
 CC Pfam; PF00008; EGF; 7.
 CC PRINTS; PR00010; EGFBL00D.
 CC PRINTS; PR00011; EGFFLAMIN.
 CC SMART; SM00051; DSL; 1.
 CC SMART; SM00179; EGF_CA; 2.
 CC PROSITE; PS00010; ASX_HYDROXYL; 1.
 CC PROSITE; PS00022; EGF_1; 8.
 CC PROSITE; PS01186; EGF_2; 7.
 CC PROSITE; PS00026; EGF_3; 8.
 CC Signal; EGF-like domain; Repeat; Transmembrane; Developmental protein;
 CC Differentiation; Glycoprotein.
 KW SIGNAL 1 26 POTENTIAL.
 KW CHAIN 27 685 DELTA-LIKE PROTEIN 4.
 FT DOMAIN 27 529 EXTRACELLULAR (POTENTIAL).
 FT TRANSMEM 530 550 POTENTIAL.
 FT DOMAIN 551 685 CYTOPLASMIC (POTENTIAL).
 FT

FT DOMAIN 155 217 DSL.
 FT DOMAIN 218 251 EGF-LIKE 1.
 FT DOMAIN 252 282 EGF-LIKE 2.
 FT DOMAIN 284 322 EGF-LIKE 3.
 FT DOMAIN 324 360 EGF-LIKE 4.
 FT DOMAIN 362 400 EGF-LIKE 5.
 FT DOMAIN 402 438 EGF-LIKE 6.
 FT DOMAIN 440 476 EGF-LIKE 7.
 FT DOMAIN 480 518 EGF-LIKE 8.
 FT DISULFID 222 233 BY SIMILARITY.
 FT DISULFID 226 239 BY SIMILARITY.
 FT DISULFID 241 250 BY SIMILARITY.
 FT DISULFID 253 254 BY SIMILARITY.
 FT DISULFID 259 270 BY SIMILARITY.
 FT DISULFID 272 281 BY SIMILARITY.
 FT DISULFID 288 300 BY SIMILARITY.
 FT DISULFID 294 310 BY SIMILARITY.
 FT DISULFID 312 321 BY SIMILARITY.
 FT DISULFID 328 339 BY SIMILARITY.
 FT DISULFID 333 348 BY SIMILARITY.
 FT DISULFID 350 359 BY SIMILARITY.
 FT DISULFID 366 377 BY SIMILARITY.
 FT DISULFID 371 388 BY SIMILARITY.
 FT DISULFID 390 399 BY SIMILARITY.
 FT DISULFID 406 417 BY SIMILARITY.
 FT DISULFID 411 426 BY SIMILARITY.
 FT DISULFID 428 437 BY SIMILARITY.
 FT DISULFID 444 464 BY SIMILARITY.
 FT DISULFID 466 475 BY SIMILARITY.
 FT DISULFID 484 495 BY SIMILARITY.
 FT DISULFID 489 506 BY SIMILARITY.
 FT DISULFID 508 517 BY SIMILARITY.
 FT CARBOHYD 108 108 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 183 183 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 205 205 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 393 393 N-LINKED (GLCNAC. .) (POTENTIAL).
 SQ SEQUENCE 685 AA; 74604 MW; 6CF89D3C220ACC89 CRC64;
 Query Match 11.4%; Score 86; DB 1; Length 685;
 Best Local Similarity 25.0%; Pred. No. 0.45;
 Matches 28; Conservative 9; Mismatches 45; Indels 30; Gaps 6;
 QY 14 WISAC-----AMLLCHGSLQHTFOOHLHPEGTCEVIAHRCCKNKRTEERSQVVKSC 69
 DB 268 WQCTCDEGWGGLFCDDQLNYC--THSPCKNGATC-----SNSGQSYT--CTC 312
 QY 70 LFGKVAGTTRNRPSCVDASIVIGRWKMEPCLEGECKTLDPNSGWMCATG 121
 DB 313 RPYGTG-----VDCELELSE--CDNPNCRNGSGCKQEDGYHCLCPPG 353
 RESULT 10
 JAGL_HUMAN
 ID JAGL_HUMAN STANDARD; PRT; 1218 AA.
 AC P78504; O14902; O15122; Q15816;
 DT 28-FEB-2003 (Rel. 41, Created)
 DT 28-FEB-2003 (Rel. 41, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Jagged 1 precursor (Jagged1) (hJ1).
 OS JAGL OR JAGL1.
 GN Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=97422615; PubMed=9268641;
 RA Oda T., Elkahoun A.G., Meltzer P.S., Chandrasekharappa S.C.;
 RT "Identification and cloning of the human homolog (JAG1) of the rat
 RT Jagged1 gene from the Alagille syndrome critical region at 20p12.";
 RL Genomics 43:376-379(1997).
 RN [2]
 RP SEQUENCE FROM N.A., AND VARIANT AGS CVS-184.
 RP

- TISSUE=Bone marrow;
 RX MEDLINE=97351506; PubMed=9207788;
 RA Li L., Krantz I.D., Deng Y., Genin A., Banta A.B., Collins C.C.,
 RA Qi M., Trask B.J., Kuo W.L., Cochran J., Costa T., Pierpont M.E.M.,
 RA Rand E.B., Piccoli D.A., Hood L., Spinner N.B.;
 RT "Alagille syndrome is caused by mutations in human Jagged1, which
 RT encodes a ligand for Notch1."
 RL Nat. Genet. 16:243-251(1997).
 RN [3]
 RP SEQUENCE FROM N.A., AND FUNCTION.
 RX MEDLINE=9812342; PubMed=9462510;
 RA Li L., Milner L.A., Deng Y., Iwata M., Banta A.B., Graf L.,
 RA Marcovina S., Friedman C., Trask B.J., Hood L., Torok-Scorb B.;
 RT "The human homolog of rat Jagged1 expressed by marrow stroma inhibits
 RT differentiation of 32D cells through interaction with Notch1."
 RL Immunity 8:43-55(1998).
 RN [4]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Cervical carcinoma;
 RX MEDLINE=99262417; PubMed=10329626;
 RA Bash J., Zong W.-X., Banga S., Rivera A., Ballard D.W., Ron Y.,
 RA Gelinac C.;
 RT "Rel/NF-kappaB can trigger the Notch signaling pathway by inducing the
 RT expression of Jagged1, a ligand for Notch receptors."
 RL EMBO J. 18:2803-2811(1999).
 RN [5]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=99180765; PubMed=10079256;
 RA Gray G.E., Mann R.S., Mitsiadis E., Henrique D., Carcangiu M.-L.,
 RA Banks A., Leiman J., Ward D., Ieh-Horowitz D., Artavanis-Teakonas S.;
 RT "Human ligands of the Notch receptor."
 RL Am. J. Pathol. 154:785-794(1999).
 RN [6]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=21638749; PubMed=11780052;
 RA Deloukas P., Matthews L.H., Ashurst J., Burton J., Gilbert J.G.R.,
 RA Jones M., Scavrides G., Almeida J.P., Babbage A.K., Baguley C.L.,
 RA Bailey J., Barlow K.F., Bates K.N., Beard L.M., Beare D.M.,
 RA Beasley O.P., Bird C.P., Blakey S.E., Bridgeman A.M., Brown A.J.,
 RA Buck D., Burrill W.D., Butler A.P., Carder C., Carter N.P.,
 RA Chapman J.C., Clamp M., Clark G., Clark L.N., Clark S.Y., Clee C.M.,
 RA Clegg S., Cobley V.E., Collier R.E., Connor R.E., Corby N.R.,
 RA Coulson A., Coville G.J., Deadman R., Dhami P.D., Dunn M.,
 RA Ellington A.G., Frankland J.A., Fraser A., French L., Garner P.,
 RA Grafham D.V., Griffiths C., Griffiths M.N.D., Gwilliam R., Hall R.E.,
 RA Hammond S., Harley J.L., Heath P.D., Ho S., Holden J.L., Howden P.J.,
 RA Huckle E., Hunt A.R., Hunt S.E., Jekosch K., Johnson C.M., Johnson D.,
 RA Lenvaeraiho M.H., Leverhwa M.A., Lloyd C., Lloyd D.M., Lovell J.D.,
 RA Marsh V.L., Martin S.B., McConachie L.J., McLay K., McMurray A.A.,
 RA Milne S.A., Mistry D., Moore M.J.F., Mullikin J.C., Nickerson T.,
 RA Oliver K., Parker A., Patel R., Pearce T.A.V., Peck A.I.,
 RA Phillimore B.J.C.T., Prathalingam S.R., Plumb R.W., Ramsay H.,
 RA Rice C.M., Ross M.T., Scott C.E., Sehra H.K., Showkeen R., Sims S.,
 RA Skuce C.D., Smith M.L., Soderlund C., Steward C.A., Sulston J.E.,
 RA Swann R.M., Sycamore N., Taylor R., Tee L., Thomas D.W., Thorpe A.,
 RA Tracey A., Tromans A.C., Vaudin M., Wall M., Wallis J.M.,
 RA Whitehead S.L., Whittaker P., Willey D.L., Williams L., Williams S.A.,
 RA Wilming L., Wray P.W., Hubbard T., Durbin R.M., Bentley D.R., Beck S.,
 RA Rogers J.;
 RT "The DNA sequence and comparative analysis of human chromosome 20."
 RL Nature 414:865-871(2001).
 RN [7]
 RP SEQUENCE OF 14-1227 FROM N.A.
 RC TISSUE=Umbilical vein endothelial cells;
 RX MEDLINE=97115768; PubMed=8955070;
 RA Zimin A.B., Pepper M.S., McMahon G.A., Nguyen F., Montesano R.,
 RA Maciag T.;
 RT "An antisense oligonucleotide to the notch ligand jagged enhances
 RT fibroblast growth factor-induced angiogenesis in vitro."
 RL J. Biol. Chem. 271:32499-32502(1996).
 RN [8]
 RP DISEASE.
 RX MEDLINE=97351505; PubMed=9207787;
 RA Oda T., Elkhoun A.G., Pike B.L., Okajima K., Krantz I.D., Genin A.,
 RA Piccoli D.A., Meltzer P.S., Spinner N.B., Collins F.S.,
 RA Chandrasekharappa S.C.;
 RT "Mutations in the human Jagged1 gene are responsible for Alagille
 RT syndrome."
 RL Nat. Genet. 16:235-242(1997).
 RN [9]
 RP DEVELOPMENTAL STAGE.
 RX MEDLINE=20436345; PubMed=10978356;
 RA Jones E.A., Clement-Jones M., Wilson D.I.;
 RT "JAGGED1 expression in human embryos: correlation with the Alagille
 RT syndrome phenotype."
 RL J. Med. Genet. 37:663-668(2000).
 RN [10]
 RP VARIANTS AGS CYS-184 AND HIS-184.
 RX MEDLINE=98254456; PubMed=9585603;
 RA Krantz I.D., Colliton R.P., Genin A., Rand E.B., Li L., Piccoli D.A.,
 RA Spinner N.B.;
 RT "Spectrum and frequency of jagged1 (JAG1) mutations in Alagille
 RT syndrome patients and their families."
 RL Am. J. Hum. Genet. 62:1361-1369(1998).
 RN [11]
 RP VARIANTS AGS HIS-79; THR-127; ARG-129; LEU-163; GLY-184; SER-187;
 RX GLY-229; PHE-284; CYS-288; PHE-438; SER-731 AND ARG-740.
 RA MEDLINE=99238888; PubMed=10220506;
 RA Crosnier C., Driancourt C., Raynaud N., Dhorne-Pollet S., Pollet N.,
 RA Bernard O., Hadchouel M., Meunier-Rotival M.;
 RT "Mutations in JAGGED1 gene are predominantly sporadic in Alagille
 RT syndrome."
 RL Gastroenterology 116:1141-1148(1999).
 RN [12]
 RP VARIANTS AGS THR-152 AND LEU-184.
 RX MEDLINE=20004539; PubMed=10533065;
 RA Pillia G., Uda M., Macis D., Frau F., Crisponi L., Balli F.,
 RA Barbera C., Colombo C., Frediani T., Gatti R., Iorio R., Marazzi M.G.,
 RA Marcellini M., Musumeci S., Nebbia G., Vajro P., Ruffa G., Zancan L.,
 RA Cao A., DeVirgili S.;
 RT "Jagged-1 mutation analysis in Italian Alagille syndrome patients."
 RL Hum. Mutat. 14:394-400(1999).
 RN [13]
 RP VARIANTS AGS TYR-229 AND ARG-386.
 RX MEDLINE=20514559; PubMed=11058998;
 RA Heritage M.L., MacMillan J.C., Colliton R.P., Genin A., Spinner N.B.,
 RA Anderson G.J.;
 RT "Jagged1 (JAG1) mutation detection in an Australian Alagille syndrome
 RT population."
 RL Hum. Mutat. 16:408-416(2000).
 RN [14]
 RP VARIANT TOF ASP-274.
 RX MEDLINE=21067871; PubMed=11152664;
 RA Eladad Z.A., Hamosh A., Biery N.J., Montgomery R.A., Duke M.,
 RA Elkins R., Dietz H.C.;
 RT "Familial Tetralogy of Fallot caused by mutation in the jagged1
 RT gene."
 RL Hum. Mol. Genet. 10:163-169(2001).
 RN [15]
 RP VARIANT AGS SER-37.
 RX MEDLINE=21096916; PubMed=11157803;
 RA Morrisette J.D., Colliton R.P., Spinner N.B.;
 RT "Defective intracellular transport and processing of JAG1 missense
 RT mutations in Alagille syndrome."
 RL Hum. Mol. Genet. 10:405-413(2001).
 RN [16]
 RP VARIANTS AGS PHE-220 AND ARG-753.
 RX MEDLINE=20579880; PubMed=11139247;
 RA Crosnier C., Driancourt C., Raynaud N., Hadchouel M.,
 RA Meunier-Rotival M.;
 RT "Fifteen novel mutations in the JAGGED1 gene of patients with Alagille
 RT syndrome."
 RL Hum. Mutat. 17:72-73(2001).
 CC -!- FUNCTION: Ligand for multiple Notch receptors and involved in the
 CC mediation of Notch signaling. May be involved in cell-fate

CC decisions during hematopoiesis. Seems to be involved in early
 CC and late stages of mammalian cardiovascular development. Inhibits
 CC myoblast differentiation (By similarity). Enhances fibroblast
 CC growth factor-induced angiogenesis (in vitro).
 CC -|- SUBUNIT: Interacts with NOTCH1, NOTCH2 AND NOTCH3 (By similarity).
 CC -|- SUBCELLULAR LOCATION: Type I membrane protein.
 CC -|- TISSUE SPECIFICITY: Widely expressed in adult and fetal tissues.
 CC In cervix epithelium expressed in undifferentiated subcolumnar
 CC reserve cells and squamous metaplasia. Expression is up-regulated
 CC in cervical squamous cell carcinoma. Expressed in bone marrow cell
 CC line HS-27a which supports the long-term maintenance of immature
 CC progenitor cells.
 CC -|- DEVELOPMENTAL STAGE: Expressed in 32-52 days embryos in the distal
 CC cardiac outflow tract and pulmonary artery, major arteries, portal
 CC vein, optic vesicle, otocyst, branchial arches, metanephros, and in
 CC pancreas, mesocardium, around the major bronchial branches, and in
 CC the neural tube.
 CC -|- DISEASE: Defects in JAG1 are the cause of Alagille syndrome (AGS)
 CC [MIM:118450]. AGS is an autosomal dominant developmental disorder
 CC that affects structures in the liver, heart, skeleton, eye,
 CC kidney and other organs.
 CC -|- DISEASE: Defects in JAG1 are a cause of tetralogy of Fallot
 CC (TOF) [MIM:187500]. TOF is a congenital heart anomaly which
 CC consists of pulmonary stenosis, ventricular septal defect,
 CC dextroposition of the aorta (aorta is on the right side instead of
 CC the left) and hypertrophy of the right ventricle. This condition
 CC results in a blue baby at birth due to inadequate oxygenation.
 CC Surgical correction is emergent.
 CC -|- SIMILARITY: Contains 15 EGF-like domains.
 CC -|- SIMILARITY: Contains 1 DSL domain.
 CC -|- CAUTION: Ref.7 sequence differs from that shown due to a
 CC frameshift in position 1187.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC
 CC Query Match 11.4%; Score 86; DB 1; Length 1218;
 CC Best Local Similarity 23.8%; Pred.No. 0.81;
 CC Matches 34; Conservative 11; Mismatches 34; Indels 64; Gaps 9;
 CC
 CC QY 37 HRP-EGTCE-----VIAHRC-----C-NKNRIERSQTVKSC 69
 CC Db 303 HQCLNGTSGNTGPDYKQSCPEGYSGPNCETAEHACLDPCHNRGSKETSIGFEC 362
 CC
 CC QY 70 LPQKVGAGTTRN-----PSCVDASIVIG-----KWW-----CEMEP 100
 CC Db 363 SPGWGTGTCSTNIDDCSPNNCSHGTTQD--LVNGFKVCVPPQWTGKTCQLDANECEAKP 420
 CC QY 101 CLSGEECKTLPDN-----SGWM 117
 CC Db 421 CVNAKCKNLIASYYCDLPGWM 443
 CC
 CC RESULT 11
 CC JAG1_MOUSE
 CC ID JAG1_MOUSE STANDARD; PRT; 1218 AA.
 CC AC Q9QXXO;
 CC DT 28-FEB-2003 (Rel. 41, Created)
 CC DT 28-FEB-2003 (Rel. 41, Last sequence update)
 CC DT 15-MAR-2004 (Rel. 43, Last annotation update)
 CC DE Jagged 1 precursor (Jagged1).
 CC GN JAG1.
 CC OS Mus musculus (Mouse).
 CC OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 CC OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 CC OX NCBI_TaxID:10090;
 CC RN [1]
 CC RP SEQUENCE FROM N.A., AND RECEPTOR INTERACTION.
 CC RC STRAIN=Swiss Webster / NIH;
 CC RX MEDLINE=20020271; PubMed=10551863;
 CC RA Shimizu K., Chiba S., Kumano K., Hosoya N., Takahashi T., Kanda Y.,
 CC RA Hamada Y., Yazaki Y., Hirai H.;
 CC RT "Mouse Jagged1 physically interacts with Notch2 and other Notch
 CC RT receptors: assessment by quantitative methods.";

RL J. Biol. Chem. 274:32961-32969 (1999).
 RL [2]
 RN SEQUENCE FROM N.A.
 RC STRAIN=C57BL/6; TISSUE=Brain;
 RX MEDLINE=22388257; PubMed=12477932;
 RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
 RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
 RA Altshul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
 RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
 RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
 RA Brownstein M.J., Udman T.B., Toshiyuki S., Carninci P., Prange C.,
 RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
 RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
 RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
 RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
 RA Fahey J., Hellon E., Kettman M., Madan A., Rodrigues S., Sanchez A.,
 RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
 RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
 RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smallos D.E.,
 RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
 RT "Generation and initial analysis of more than 15,000 full-length
 RL human and mouse cDNA sequences.";
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).
 RN [3]
 RP TISSUE SPECIFICITY.
 RX MEDLINE=20025753; PubMed=10556292;
 RA Loomes K.M., Underkofer L.A., Morabito J., Gottlieb S.,
 RA Piccoli D.A., Spinner N.B., Baldwin H.S., Oakley R.J.;
 RT "The expression of Jagged1 in the developing mammalian heart
 RT correlates with cardiovascular disease in Alagille syndrome.";
 RL Hum. Mol. Genet. 8:2443-2449 (1999).
 RL CC -|- FUNCTION: Ligand for multiple Notch receptors and involved in the
 CC mediation of Notch signaling. May be involved in cell-fate
 CC decisions during hematopoiesis. Seems to be involved in early
 CC and late stages of mammalian cardiovascular development. Inhibits
 CC myoblast differentiation (By similarity). May regulate fibroblast
 CC growth factor-induced angiogenesis.
 CC -|- SUBUNIT: Interacts with NOTCH1, NOTCH2 AND NOTCH3.
 CC -|- SUBCELLULAR LOCATION: Type I membrane protein.
 CC -|- TISSUE SPECIFICITY: Widely expressed in many tissues, with highest
 CC expression in brain, heart, muscle and thymus.
 CC -|- DEVELOPMENTAL STAGE: At 8.75-9.75 dpc expression was detected
 CC in structures that include those destined to contribute to the
 CC cardiovascular system of the adult heart. Expression was also
 CC detected in the mesencephalon and rhombencephalon.
 CC -|- DOMAIN: The DSL domain is indispensable and sufficient for binding
 CC to NOTCH2.
 CC -|- SIMILARITY: Contains 15 EGF-like domains.
 CC -|- SIMILARITY: Contains 1 DSL domain.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; AF171092; AAF15505.1; -;
 CC EMBL; BC058675; AAH58675.1; -;
 CC HSP; P00740; 1EDM.
 CC MGD; MGI:1095416; Jag1.
 CC GO; GO:0005576; C:extracellular; ISS.
 CC GO; GO:0005509; F:calcium ion binding; NAS.
 CC GO; GO:0008083; F:growth factor activity; ISS.
 CC GO; GO:0005112; F:Notch binding; IPI.
 CC GO; GO:0005198; F:structural molecule activity; ISS.
 CC GO; GO:0001525; P:angiogenesis; ISS.
 CC GO; GO:0001709; P:cell fate determination; ISS.
 CC GO; GO:0045446; P:endothelial cell differentiation; ISS.
 CC GO; GO:0030097; P:hematopoiesis; ISS.

DR GO:00020216; P:keratinocyte differentiation; ISS.
 DR GO:00020111; P:morphogenesis of an epithelial sheet; IMP.
 DR GO:0045445; P:myoblast differentiation; ISS.
 DR GO:0007219; P:N signaling pathway; ISS.
 DR GO:0007399; P:neurogenesis; ISS.
 DR GO:0042127; P:regulation of cell proliferation; ISS.
 DR InterPro: IPR000152; Asx_hydroxyl_S.
 DR InterPro: IPR001774; DSL.
 DR InterPro: IPR000742; EGF_2.
 DR InterPro: IPR001881; EGF_CA.
 DR InterPro: IPR001438; EGF_II.
 DR InterPro: IPR006209; EGF_Like.
 DR InterPro: IPR001007; VWF_C.
 DR Pfam: PF01444; DSL; 1.
 DR Pfam: PF00008; EGF; 14.
 DR PRINTS: PR00010; EGFBLD.
 DR SMART: SMO0051; DSL; 1.
 DR SMART: SMO0179; EGF_CA; 10.
 DR SMART: SMO0214; VWC; 1.
 DR PROSITE: PS00010; ASX_HYDROXYL; 10.
 DR PROSITE: PS00022; EGF_1; 16.
 DR PROSITE: PS01186; EGF_2; 12.
 DR PROSITE: PS00026; EGF_3; 15.
 DR PROSITE: PS01187; EGF_CA; 8.
 DR Calcium-binding; EGF-Like domain; Glycoprotein; Developmental protein;
 KW Repeat; Transmembrane; Signal.
 FT SIGNAL 1 33 POTENTIAL.
 FT CHAIN 34 1218 JAGGED 1.
 FT DOMAIN 34 1067 EXTRACELLULAR (POTENTIAL).
 FT TRANSMEM 1068 1093 POTENTIAL.
 FT DOMAIN 1094 1218 CYTOPLASMIC (POTENTIAL).
 FT DOMAIN 167 229 DSL.
 FT DOMAIN 230 266 EGF-LIKE 1 (ATYPICAL).
 FT DOMAIN 296 334 EGF-LIKE 2.
 FT DOMAIN 336 372 EGF-LIKE 3.
 FT DOMAIN 374 410 EGF-LIKE 4, CALCIUM-BINDING (POTENTIAL).
 FT DOMAIN 412 448 EGF-LIKE 5, CALCIUM-BINDING (POTENTIAL).
 FT DOMAIN 450 485 EGF-LIKE 6, CALCIUM-BINDING (POTENTIAL).
 FT DOMAIN 487 523 EGF-LIKE 7, CALCIUM-BINDING (POTENTIAL).
 FT DOMAIN 525 561 EGF-LIKE 8.
 FT DOMAIN 574 627 EGF-LIKE 9.
 FT DOMAIN 629 665 EGF-LIKE 10, CALCIUM-BINDING (POTENTIAL).
 FT DOMAIN 667 703 EGF-LIKE 11, CALCIUM-BINDING (POTENTIAL).
 FT DOMAIN 705 741 EGF-LIKE 12.
 FT DOMAIN 744 780 EGF-LIKE 13.
 FT DOMAIN 782 818 EGF-LIKE 14, CALCIUM-BINDING (POTENTIAL).
 FT DOMAIN 820 856 EGF-LIKE 15, CALCIUM-BINDING (POTENTIAL).
 FT DISULFID 300 312 BY SIMILARITY.
 FT DISULFID 306 322 BY SIMILARITY.
 FT DISULFID 324 333 BY SIMILARITY.
 FT DISULFID 340 351 BY SIMILARITY.
 FT DISULFID 345 360 BY SIMILARITY.
 FT DISULFID 362 371 BY SIMILARITY.
 FT DISULFID 378 389 BY SIMILARITY.
 FT DISULFID 383 398 BY SIMILARITY.
 FT DISULFID 400 409 BY SIMILARITY.
 FT DISULFID 416 427 BY SIMILARITY.
 FT DISULFID 421 436 BY SIMILARITY.
 FT DISULFID 438 447 BY SIMILARITY.
 FT DISULFID 454 464 BY SIMILARITY.
 FT DISULFID 458 473 BY SIMILARITY.
 FT DISULFID 475 484 BY SIMILARITY.
 FT DISULFID 491 502 BY SIMILARITY.
 FT DISULFID 496 511 BY SIMILARITY.
 FT DISULFID 513 522 BY SIMILARITY.
 FT DISULFID 529 540 BY SIMILARITY.
 FT DISULFID 534 549 BY SIMILARITY.
 FT DISULFID 551 560 BY SIMILARITY.
 FT DISULFID 551 560 BY SIMILARITY.
 FT DISULFID 578 605 BY SIMILARITY.
 FT DISULFID 599 615 BY SIMILARITY.
 FT DISULFID 617 626 BY SIMILARITY.
 FT DISULFID 633 644 BY SIMILARITY.
 FT DISULFID 638 653 BY SIMILARITY.

FT DISULFID 655 664 BY SIMILARITY.
 FT DISULFID 671 682 BY SIMILARITY.
 FT DISULFID 676 691 BY SIMILARITY.
 FT DISULFID 693 702 BY SIMILARITY.
 FT DISULFID 709 720 BY SIMILARITY.
 FT DISULFID 714 729 BY SIMILARITY.
 FT DISULFID 731 740 BY SIMILARITY.
 FT DISULFID 748 759 BY SIMILARITY.
 FT DISULFID 753 768 BY SIMILARITY.
 FT DISULFID 770 779 BY SIMILARITY.
 FT DISULFID 786 797 BY SIMILARITY.
 FT DISULFID 791 806 BY SIMILARITY.
 FT DISULFID 808 817 BY SIMILARITY.
 FT DISULFID 824 835 BY SIMILARITY.
 FT DISULFID 829 844 BY SIMILARITY.
 FT DISULFID 846 855 BY SIMILARITY.
 FT CARBOHYD 143 143 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 217 217 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 382 382 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 559 559 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 745 745 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 960 960 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 991 991 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 1045 1045 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 1064 1064 N-LINKED (GLCNAC. .) (POTENTIAL).
 SQ SEQUENCE 1218 AA; 134163 MW; 77739F8928BB793C CRC64;
 Query Match 11.4%; Score 86; DB 1; Length 1218;
 Best Local Similarity 27.0%; Pred. No. 0.81;
 Matches 33; Conservative 9; Mismatches 34; Indels 46; Gaps 9;
 QY 39 PE---GTCCEVIAAHC-----C-NKRIERSQTVKSCLPKGVAGTTRNR----- 81
 DB 325 PEGYSGPNC-IAEHACLSDPCHNRGCKTSGFCECSFGTGTCTSNIDDCSPNNC 383
 QY 82 ---PSCVDASIVIG-----KWW-----CEMEPCLEGECKTLDPN-----SG 115
 DB 384 SHGCTQD---LVNGFKVCVCPQWTGTCQLDANECAKPCVNARSCNLIASYCDCLPG 441
 QY 116 WM 117
 DB 442 WM 443
 RESULT 12
 JAG1_RAT STANDARD; PRT; 1219 AA.
 ID JAG1_RAT
 AC Q63722; P70640;
 DT 28-FEB-2003 (Rel. 41, Created)
 DT 28-FEB-2003 (Rel. 41, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Jagged 1 precursor (Jaggedl).
 OS Rattus norvegicus (Rat).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
 OX NCBI_TaxID=10116;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Sciatic nerve;
 RX MEDLINE=95211842; PubMed=7697721;
 RA Lindeell C.E., Shawber C.J., Boulter J., Weinmaster G.;
 RT "Jagged: a mammalian ligand that activates Notch1.";
 RL Cell 80:909-917(1995).
 CC -!- FUNCTION: Ligand for multiple Notch receptors and involved in the
 CC mediation of Notch signaling. May be involved in cell-fate
 CC decisions during hematopoiesis. Enhances fibroblast growth
 CC factor-induced angiogenesis (in vitro). Seems to be involved in
 CC early and late stages of mammalian cardiovascular development.
 CC Inhibits myoblast differentiation. May regulate fibroblast
 CC growth factor-induced angiogenesis.
 CC -!- SUBUNIT: Interacts with NOTCH1, NOTCH2 AND NOTCH3 (By similarity).
 CC -!- SUBCELLULAR LOCATION: Type I membrane protein.
 CC -!- TISSUE SPECIFICITY: Widely expressed in a variety of tissues.

FT	DISULFID	324	333	BY SIMILARITY.
FT	DISULFID	340	351	BY SIMILARITY.
FT	DISULFID	345	360	BY SIMILARITY.
FT	DISULFID	362	371	BY SIMILARITY.
FT	DISULFID	378	389	BY SIMILARITY.
FT	DISULFID	383	398	BY SIMILARITY.
FT	DISULFID	400	409	BY SIMILARITY.
FT	DISULFID	416	427	BY SIMILARITY.
FT	DISULFID	421	436	BY SIMILARITY.
FT	DISULFID	438	447	BY SIMILARITY.
FT	DISULFID	454	464	BY SIMILARITY.
FT	DISULFID	458	473	BY SIMILARITY.
FT	DISULFID	475	484	BY SIMILARITY.
FT	DISULFID	491	502	BY SIMILARITY.
FT	DISULFID	496	511	BY SIMILARITY.
FT	DISULFID	513	522	BY SIMILARITY.
FT	DISULFID	529	540	BY SIMILARITY.
FT	DISULFID	534	549	BY SIMILARITY.
FT	DISULFID	551	560	BY SIMILARITY.
FT	DISULFID	578	605	BY SIMILARITY.
FT	DISULFID	599	615	BY SIMILARITY.
FT	DISULFID	617	626	BY SIMILARITY.
FT	DISULFID	633	644	BY SIMILARITY.
FT	DISULFID	638	653	BY SIMILARITY.
FT	DISULFID	655	664	BY SIMILARITY.
FT	DISULFID	671	682	BY SIMILARITY.
FT	DISULFID	676	691	BY SIMILARITY.
FT	DISULFID	693	702	BY SIMILARITY.
FT	DISULFID	709	720	BY SIMILARITY.
FT	DISULFID	714	729	BY SIMILARITY.
FT	DISULFID	731	740	BY SIMILARITY.
FT	DISULFID	748	759	BY SIMILARITY.
FT	DISULFID	753	768	BY SIMILARITY.
FT	DISULFID	770	779	BY SIMILARITY.
FT	DISULFID	786	797	BY SIMILARITY.
FT	DISULFID	791	806	BY SIMILARITY.
FT	DISULFID	808	817	BY SIMILARITY.
FT	DISULFID	824	835	BY SIMILARITY.
FT	DISULFID	829	844	BY SIMILARITY.
FT	DISULFID	846	855	BY SIMILARITY.
FT	CARBOHYD	143	143	N-LINKED (GLCNAC. .) (POTENTIAL).
FT	CARBOHYD	217	217	N-LINKED (GLCNAC. .) (POTENTIAL).
FT	CARBOHYD	382	382	N-LINKED (GLCNAC. .) (POTENTIAL).
FT	CARBOHYD	559	559	N-LINKED (GLCNAC. .) (POTENTIAL).
FT	CARBOHYD	745	745	N-LINKED (GLCNAC. .) (POTENTIAL).
FT	CARBOHYD	960	960	N-LINKED (GLCNAC. .) (POTENTIAL).
FT	CARBOHYD	991	991	N-LINKED (GLCNAC. .) (POTENTIAL).
FT	CARBOHYD	1045	1045	N-LINKED (GLCNAC. .) (POTENTIAL).
FT	CARBOHYD	1064	1064	N-LINKED (GLCNAC. .) (POTENTIAL).
FT	SEQUENCE	1219 AA; 134325 MW; 65D4CFC238A0E204 CRC64;		
Query Match		11.4%;	Score 86; DB 1; Length 1219;	
Best Local Similarity		27.0%;	Pred. No. 0.81;	
Matches 33; Conservative		9;	Mismatches 34; Indels 46; Gaps 9;	
QY	39	PE---GGTCEVIAHRC	---C-NKNRIERSQTVKSCLPKGVAGTRNR	----- 81
Db	325	PEGYSGNCE-IAEHACLSDPCHNRSGKSTSGFCECECPGWTGPTCTSNIDDCSPNC	383	
QY	82	---PSCVDASIVIG	-----KWW-----CEMEPCLEGECKTLFDPN	-----SG 115
Db	384	SHGTTQCD--LVNGFKVCVPQWTGKTQCLDANECAKPCVNARSKNLIASYVCDCLPG	441	
QY	116	WM 117		
Db	442	WM 443		
RESULT 13				
FBP1_STRPU				
ID	FBP1_STRPU	STANDARD;	PRT;	1064 AA.
AC	P10079;			
DT	01-MAR-1989 (Rel. 10, Created)			

DT 01-FEB-1996 (Rel. 33, Last sequence update)
 DT 15-MAR-2004 (Rel. 43, Last annotation update)
 DE Fibropellin I precursor (Epidermal growth factor-related protein 1)
 DE (UEGF-1).
 GN EGF1.
 OS Strongylocentrotus purpuratus (Purple sea urchin).
 OC Eukaryota; Metazoa; Echinodermata; Eleutherozoa; Echinozoa;
 OC Echinoidea; Euechinozoa; Echinacea; Echinoidea; Strongylocentrotidae;
 OC Strongylocentrotus.
 OX NCBI_TaxID=7668;
 RN [1]
 RN SEQUENCE FROM N.A.
 RP MEDLINE=90112459; PubMed=2514273;
 RA Delgadillo-Reynoso M.G., Rollo D.R., Hursh D.A., Raff R.A.;
 RT "Structural analysis of the UEGF gene in the sea urchin
 RT strongylocentrotus purpuratus reveals more similarity to vertebrate
 RT than to invertebrate genes with EGF-like repeats.";
 RL J. Mol. Evol. 29:314-327(1989).
 RN [2]
 RN SEQUENCE OF 279-476 AND 791-1064 FROM N.A.
 RP MEDLINE=87319677; PubMed=3498216;
 RA Hursh D.A., Andrews M.E., Raff R.A.;
 RT "A sea urchin gene encodes a polypeptide homologous to epidermal
 RT growth factor.";
 RL Science 237:1487-1490(1987).
 RN [3]
 RN AVIDIN-LIKE DOMAIN.
 RP MEDLINE=89196806; PubMed=2784773;
 RA Hunt L.T., Barker W.C.;
 RT "Avidin-like domain in an epidermal growth factor homolog from a sea
 RT urchin.";
 RL FASEB J. 3:1760-1764(1989).
 RN [4]
 RN CHARACTERIZATION.
 RP MEDLINE=91285254; PubMed=2060714;
 RA Bisgrove B.W., Andrews M.E., Raff R.A.;
 RT "Fibropellins, products of an EGF repeat-containing gene, form a
 RT unique extracellular matrix structure that surrounds the sea urchin
 RT embryo.";
 RL Dev. Biol. 146:89-99(1991).
 CC -|- FUNCTION: Form the apical lamina, a component of the extracellular
 CC matrix.
 CC -|- SUBCELLULAR LOCATION: EXTRACELLULAR, IN VESICLES IN THE CYTOPLASM
 CC OF UNFERTILIZED EGGS, THEN TO THE BASE OF THE HYALIN LAYER
 CC THROUGHOUT DEVELOPMENT AND FINALLY IN THE APICAL LAMINA IN LATE
 CC EMBRYOS AND EARLY LARVAE.
 CC -|- ALTERNATIVE PRODUCTS:
 CC Event=Alternative splicing; Named isoforms=2;
 CC Name=IA;
 CC IsoId=PI0079-1; Sequence=Displayed;
 CC Name=IB;
 CC IsoId=PI0079-2; Sequence=VSP_000451;
 CC -|- DEVELOPMENTAL STAGE: Moderate levels in unfertilized eggs and
 CC during early cleavage, then rapidly increases in abundance between
 CC late morula and mesenchyme blastula stages to maximal levels
 CC maintained through subsequent stages. Expressed both maternally
 CC and zygotically.
 CC -|- SIMILARITY: Contains 21 EGF-like domains.
 CC -|- SIMILARITY: Contains 1 CUB domain.
 CC -|- SIMILARITY: THE C-TERMINAL DOMAIN OF THIS PROTEIN IS SIMILAR TO
 CC AVIDIN/SREPTAVIDIN.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; L08692; AAA62164.1; -
 CC EMBL; L08692; AAA62163.1; -
 CC EMBL; X17530; CAA35571.1; -

DR EMBL; M17421; AAA30050.1; -
 DR EMBL; X17533; CAA35573.1; -
 DR PIR; A40136; A40136.
 DR HSSP; P01132; 1EGF.
 DR InterPro; IPR000152; Asx_hydroxyl_s.
 DR InterPro; IPR005469; Avidin.
 DR InterPro; IPR005468; Avidin/str.
 DR InterPro; IPR000859; CUB.
 DR InterPro; IPR000742; EGF_2.
 DR InterPro; IPR001881; EGF_Ca.
 DR InterPro; IPR001438; EGF_II.
 DR InterPro; IPR006209; EGF_Like.
 DR Pfam; PF01382; Avidin; 1.
 DR Pfam; PF00431; CUB; 1.
 DR Pfam; PF00008; EGF; 21.
 DR PRINTS; PR00709; AVIDIN.
 DR PRINTS; PR00010; EGFLOOD.
 DR SMART; SM00042; CUB; 1.
 DR SMART; SM00179; EGF_Ca; 20.
 DR PROSITE; PS00010; ASX_HYDROXYL; 19.
 DR PROSITE; PS00577; AVIDIN; 1.
 DR PROSITE; PS01180; CUB; 1.
 DR PROSITE; PS00022; EGF_1; 19.
 DR PROSITE; PS01186; EGF_2; 19.
 DR PROSITE; PS50026; EGF_3; 21.
 DR PROSITE; PS01187; EGF_Ca; 18.
 KW Biotin; Alternative splicing; EGF-like domain; Repeat; Signal;
 KW Glycoprotein; Calcium-binding.
 FT SIGNAL 1 19 POTENTIAL.
 FT CHAIN 20 1064 FIBROPELLIN 1.
 FT DOMAIN 20 55 EGF-LIKE 1.
 FT DOMAIN 62 175 CUB.
 FT DOMAIN 176 212 EGF-LIKE 2, CALCIUM-BINDING (POTENTIAL).
 FT DOMAIN 214 250 EGF-LIKE 3, CALCIUM-BINDING (POTENTIAL).
 FT DOMAIN 252 288 EGF-LIKE 4, CALCIUM-BINDING (POTENTIAL).
 FT DOMAIN 290 326 EGF-LIKE 5, CALCIUM-BINDING (POTENTIAL).
 FT DOMAIN 328 364 EGF-LIKE 6, CALCIUM-BINDING (POTENTIAL).
 FT DOMAIN 366 402 EGF-LIKE 7, CALCIUM-BINDING (POTENTIAL).
 FT DOMAIN 404 440 EGF-LIKE 8, CALCIUM-BINDING (POTENTIAL).
 FT DOMAIN 442 478 EGF-LIKE 9, CALCIUM-BINDING (POTENTIAL).
 FT DOMAIN 480 516 EGF-LIKE 10, CALCIUM-BINDING (POTENTIAL).
 FT DOMAIN 518 556 EGF-LIKE 11, CALCIUM-BINDING (POTENTIAL).
 FT DOMAIN 556 592 EGF-LIKE 12, CALCIUM-BINDING (POTENTIAL).
 FT DOMAIN 594 630 EGF-LIKE 13, CALCIUM-BINDING (POTENTIAL).
 FT DOMAIN 632 668 EGF-LIKE 14, CALCIUM-BINDING (POTENTIAL).
 FT DOMAIN 670 706 EGF-LIKE 15, CALCIUM-BINDING (POTENTIAL).
 FT DOMAIN 708 744 EGF-LIKE 16, CALCIUM-BINDING (POTENTIAL).
 FT DOMAIN 746 782 EGF-LIKE 17, CALCIUM-BINDING (POTENTIAL).
 FT DOMAIN 784 820 EGF-LIKE 18, CALCIUM-BINDING (POTENTIAL).
 FT DOMAIN 822 858 EGF-LIKE 19, CALCIUM-BINDING (POTENTIAL).
 FT DOMAIN 860 896 EGF-LIKE 20, CALCIUM-BINDING (POTENTIAL).
 FT DOMAIN 898 934 EGF-LIKE 21, CALCIUM-BINDING (POTENTIAL).
 FT DOMAIN 936 1064 AVIDIN-LIKE.
 FT DISULFID 23 34 BY SIMILARITY.
 FT DISULFID 28 43 BY SIMILARITY.
 FT DISULFID 45 54 BY SIMILARITY.
 FT DISULFID 180 191 BY SIMILARITY.
 FT DISULFID 185 200 BY SIMILARITY.
 FT DISULFID 202 211 BY SIMILARITY.
 FT DISULFID 218 229 BY SIMILARITY.
 FT DISULFID 223 238 BY SIMILARITY.
 FT DISULFID 240 249 BY SIMILARITY.
 FT DISULFID 256 267 BY SIMILARITY.
 FT DISULFID 261 276 BY SIMILARITY.
 FT DISULFID 278 287 BY SIMILARITY.
 FT DISULFID 294 305 BY SIMILARITY.
 FT DISULFID 299 314 BY SIMILARITY.
 FT DISULFID 316 325 BY SIMILARITY.
 FT DISULFID 332 343 BY SIMILARITY.
 FT DISULFID 337 352 BY SIMILARITY.
 FT DISULFID 354 363 BY SIMILARITY.
 FT DISULFID 370 381 BY SIMILARITY.
 FT DISULFID 375 390 BY SIMILARITY.

FT	DISULFID	392	401	BY SIMILARITY.
FT	DISULFID	408	419	BY SIMILARITY.
FT	DISULFID	413	428	BY SIMILARITY.
FT	DISULFID	430	439	BY SIMILARITY.
FT	DISULFID	446	457	BY SIMILARITY.
FT	DISULFID	451	466	BY SIMILARITY.
FT	DISULFID	468	477	BY SIMILARITY.
FT	DISULFID	484	495	BY SIMILARITY.
FT	DISULFID	489	504	BY SIMILARITY.
FT	DISULFID	506	515	BY SIMILARITY.
FT	DISULFID	522	533	BY SIMILARITY.
FT	DISULFID	527	542	BY SIMILARITY.
FT	DISULFID	544	553	BY SIMILARITY.
FT	DISULFID	560	571	BY SIMILARITY.
FT	DISULFID	565	580	BY SIMILARITY.
FT	DISULFID	582	591	BY SIMILARITY.
FT	DISULFID	598	609	BY SIMILARITY.
FT	DISULFID	603	618	BY SIMILARITY.
FT	DISULFID	620	629	BY SIMILARITY.
FT	DISULFID	636	647	BY SIMILARITY.
FT	DISULFID	641	656	BY SIMILARITY.
FT	DISULFID	658	667	BY SIMILARITY.
FT	DISULFID	674	685	BY SIMILARITY.
FT	DISULFID	679	694	BY SIMILARITY.
FT	DISULFID	696	705	BY SIMILARITY.
FT	DISULFID	712	723	BY SIMILARITY.
FT	DISULFID	717	732	BY SIMILARITY.
FT	DISULFID	734	743	BY SIMILARITY.
FT	DISULFID	750	761	BY SIMILARITY.
FT	DISULFID	755	770	BY SIMILARITY.
FT	DISULFID	772	781	BY SIMILARITY.
FT	DISULFID	788	799	BY SIMILARITY.
FT	DISULFID	793	808	BY SIMILARITY.
FT	DISULFID	810	819	BY SIMILARITY.
FT	DISULFID	826	837	BY SIMILARITY.
FT	DISULFID	831	846	BY SIMILARITY.
FT	DISULFID	848	857	BY SIMILARITY.
FT	DISULFID	864	875	BY SIMILARITY.
FT	DISULFID	869	884	BY SIMILARITY.
FT	DISULFID	886	895	BY SIMILARITY.
FT	DISULFID	902	913	BY SIMILARITY.
FT	DISULFID	907	922	BY SIMILARITY.
FT	DISULFID	924	933	BY SIMILARITY.
FT	CARBOHYD	30	30	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	136	136	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	851	851	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	VARSPLOC	477	780	Missing (in isoform IB).
FT	CONFLICT	279	279	L -> S (IN REF. 2).
FT	SEQUENCE	1064 AA;	112072 MW;	2E569CA012ED6D09 CRC64;
Query Match				
Best Local Similarity 11.3%; Score 85.5; DB 1; Length 1064;				
Matches 25; Conservative 9; Mismatches 34; Indels 21; Gaps 4;				
QY	40	EGGTCEV----	IAAHRCKNKNRIERSQTVKSCILPGKV----	AGTTRNPSCVDASIVIG 92
Db	625	EGTNCETINTDCASPCNMNGGLCVQVNSYVCFCLPGFTGTHCGTE-----	IDR-----	673
QY	93	KWCMEPECLEGECKTLPDNSGWMCAWG 121		
Db	674	---CASSFLNGGQCIDRVDVSVECVCAAG 699		
RESULT 14				
NTC4 HUMAN				
AC	ID	NTC4_HUMAN	STANDARD;	PRT: 2003 AA.
AD	Q99456;	Q00306;	Q99458;	Q9H358; Q9UIJ0;
DT	28-FEB-2003	(Rel. 41, Created)		
DT	28-FEB-2003	(Rel. 41, Last sequence update)		
DT	28-FEB-2003	(Rel. 41, Last annotation update)		
DE	Neurogenic locus notch homolog protein 4 precursor (Notch 4) (hNotch4).			

FT	DISULFID	392	401	BY SIMILARITY.
FT	DISULFID	408	419	BY SIMILARITY.
FT	DISULFID	413	428	BY SIMILARITY.
FT	DISULFID	430	439	BY SIMILARITY.
FT	DISULFID	446	457	BY SIMILARITY.
FT	DISULFID	451	466	BY SIMILARITY.
FT	DISULFID	468	477	BY SIMILARITY.
FT	DISULFID	484	495	BY SIMILARITY.
FT	DISULFID	489	504	BY SIMILARITY.
FT	DISULFID	506	515	BY SIMILARITY.
FT	DISULFID	522	533	BY SIMILARITY.
FT	DISULFID	527	542	BY SIMILARITY.
FT	DISULFID	544	553	BY SIMILARITY.
FT	DISULFID	560	571	BY SIMILARITY.
FT	DISULFID	565	580	BY SIMILARITY.
FT	DISULFID	582	591	BY SIMILARITY.
FT	DISULFID	598	609	BY SIMILARITY.
FT	DISULFID	603	618	BY SIMILARITY.
FT	DISULFID	620	629	BY SIMILARITY.
FT	DISULFID	636	647	BY SIMILARITY.
FT	DISULFID	641	656	BY SIMILARITY.
FT	DISULFID	658	667	BY SIMILARITY.
FT	DISULFID	674	685	BY SIMILARITY.
FT	DISULFID	679	694	BY SIMILARITY.
FT	DISULFID	696	705	BY SIMILARITY.
FT	DISULFID	712	723	BY SIMILARITY.
FT	DISULFID	717	732	BY SIMILARITY.
FT	DISULFID	734	743	BY SIMILARITY.
FT	DISULFID	750	761	BY SIMILARITY.
FT	DISULFID	755	770	BY SIMILARITY.
FT	DISULFID	772	781	BY SIMILARITY.
FT	DISULFID	788	799	BY SIMILARITY.
FT	DISULFID	793	808	BY SIMILARITY.
FT	DISULFID	810	819	BY SIMILARITY.
FT	DISULFID	826	837	BY SIMILARITY.
FT	DISULFID	831	846	BY SIMILARITY.
FT	DISULFID	848	857	BY SIMILARITY.
FT	DISULFID	864	875	BY SIMILARITY.
FT	DISULFID	869	884	BY SIMILARITY.
FT	DISULFID	886	895	BY SIMILARITY.
FT	DISULFID	902	913	BY SIMILARITY.
FT	DISULFID	907	922	BY SIMILARITY.
FT	DISULFID	924	933	BY SIMILARITY.
FT	CARBOHYD	30	30	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	136	136	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	851	851	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	VARSPLOC	477	780	Missing (in isoform IB).
FT	CONFLICT	279	279	L -> S (IN REF. 2).
FT	SEQUENCE	1064 AA;	112072 MW;	2E569CA012ED6D09 CRC64;
Query Match		11.3%;	Score 85.5;	DB 1; Length 1064;
Best Local Similarity		28.1%;	Pred. No. 0.79;	
Matches		25;	Conservative	9; Mismatches 34; Indels 21; Gaps 4;
QY	40	EGGTCEV----	IAAHRCKNKNRIERSQTVKSCILPGKV----	AGTTRNPSCVDASIVIG 92
Db	625	EGTWCEINTD	CASPCNMNGGLCVQVNSYVCFCLPGFTGTHCGTE-----	IDF----- 673
QY	93	KWCEMEPCLEGECKTLPD	NSGWMCAWG 121	
Db	674	---CASSFLNGGQCID	RDVDSVEVCVAAG 699	
RESULT 14				
NTC4 HUMAN				
ID	NTC4_HUMAN	STANDARD;	PRT: 2003 AA.	
AC	Q99456;	Q00306;	Q99458;	Q9H358; Q9UIJ0;
DT	28-FEB-2003	(Rel. 41, Created)		
DT	28-FEB-2003	(Rel. 41, Last sequence update)		
DT	28-FEB-2003	(Rel. 41, Last annotation update)		
DE	Neurogenic locus notch homolog protein 4 precursor (Notch 4) (hNtch4).			

from the membrane (By similarity).
 -!- PTM: Phosphorylated (By similarity).
 -!- POLYMORPHISM: The poly-Leu region of NOTCH4 (in the signal peptide) is polymorphic and the number of Leu varies in the population (from 6 to 12).
 -!- SIMILARITY: Belongs to the NOTCH family.
 -!- SIMILARITY: Contains 28 EGF-like domains.
 -!- SIMILARITY: Contains 3 lin/Notch repeats.
 -!- SIMILARITY: Contains 5 ANK repeats.
 -!- CAUTION: Ref.1 sequence differs from that shown due to frameshifts in position 1438 to 1463.

 This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

 EMBL; D63395; BAA09708.1; ALT_FRAME.
 EMBL; D6566; BAA13116.1; -.
 EMBL; U95299; AAC32288.1; -.
 EMBL; U93335; AAC63097.1; -.
 EMBL; AB023961; BAB20317.1; -.
 EMBL; AB024520; BAA88951.1; -.
 EMBL; AB024578; BAA88952.1; -.
 HSSP; P08709; 1BF9.
 Genew; HGNC:7884; NOTCH4.
 MIM; 164951; -.
 InterPro; IPR002110; ANK.
 InterPro; IPR000152; ASX_hydroxyl_S.
 InterPro; IPR000742; EGF_2.
 InterPro; IPR001881; EGF_CA.
 InterPro; IPR001438; EGF_II.
 InterPro; IPR006209; EGF_like.
 InterPro; IPR002049; Laminin_EGF.
 InterPro; IPR008297; Notch.
 InterPro; IPR000800; Notch_dom.
 Pfam; PF00023; ank; 6.
 Pfam; PF00066; notch; 2.
 Pfam; PF00008; EGF; 26.
 PIRSF; PIRSF02279; Notch; 1.
 PRINTS; PR00010; EGFLOOD.
 PRINTS; PR00011; EGF_LAMININ.
 PRINTS; PR01452; NOTCH.
 SMART; SM00248; ANK; 5.
 SMART; SM00179; EGF_CA; 11.
 SMART; SM00004; NL; 3.
 PROSITE; PS0297; ANK_REPEAT_REGION; 1.
 PROSITE; PS0088; ANK_REPEAT; 5.
 PROSITE; PS0010; ASX_HYDROXYL; 11.
 PROSITE; PS00022; EGF_1; 28.
 PROSITE; PS01186; EGF_2; 21.
 PROSITE; PS00026; EGF_3; 28.
 PROSITE; PS01187; EGF_CA; 9.
 Receptor; transcription regulation; Activator; Differentiation;
 Developmental protein; Repeat; ANK repeat; EGF-like domain;
 Transmembrane; Glycoprotein; Signal; Phosphorylation; Polymorphism;
 Triplet repeat expansion; Alternative splicing.
 SIGNAL 1 23 POTENTIAL
 CHAIN 24 2003 NEUROGENIC LOCUS NOTCH PROTEIN HOMOLOG 4.
 CHAIN 1432 2003 NOTCH EXTRACELLULAR TRUNCATION
 (BY SIMILARITY).
 CHAIN 1467 2003 NOTCH INTRACELLULAR DOMAIN
 (BY SIMILARITY).
 DOMAIN 24 1447 EXTRACELLULAR (POTENTIAL).
 TRANSMEM 1448 1468 POTENTIAL.
 DOMAIN 1469 2003 CYTOPLASMIC (POTENTIAL).
 DOMAIN 24 63 EGF-LIKE 1.
 DOMAIN 64 115 EGF-LIKE 2.
 DOMAIN 118 155 EGF-LIKE 3.
 DOMAIN 156 192 EGF-LIKE 4.

FT	DOMAIN	194	232	EGF-LIKE 5, CALCIUM-BINDING (POTENTIAL).
FT	DOMAIN	234	274	EGF-LIKE 6.
FT	DOMAIN	276	312	CALCIUM-BINDING (POTENTIAL).
FT	DOMAIN	314	353	EGF-LIKE 8, CALCIUM-BINDING (POTENTIAL).
FT	DOMAIN	355	391	EGF-LIKE 9, CALCIUM-BINDING (POTENTIAL).
FT	DOMAIN	392	430	EGF-LIKE 10.
FT	DOMAIN	432	473	EGF-LIKE 11.
FT	DOMAIN	475	511	EGF-LIKE 12, CALCIUM-BINDING (POTENTIAL).
FT	DOMAIN	513	549	EGF-LIKE 13, CALCIUM-BINDING (POTENTIAL).
FT	DOMAIN	551	587	EGF-LIKE 14, CALCIUM-BINDING (POTENTIAL).
FT	DOMAIN	589	625	EGF-LIKE 15, CALCIUM-BINDING (POTENTIAL).
FT	DOMAIN	626	659	EGF-LIKE 16.
FT	DOMAIN	661	689	EGF-LIKE 17.
FT	DOMAIN	691	727	EGF-LIKE 18.
FT	DOMAIN	729	765	EGF-LIKE 19.
FT	DOMAIN	767	803	EGF-LIKE 20.
FT	DOMAIN	806	842	EGF-LIKE 21.
FT	DOMAIN	844	880	EGF-LIKE 22.
FT	DOMAIN	882	928	EGF-LIKE 23.
FT	DOMAIN	930	966	EGF-LIKE 24.
FT	DOMAIN	968	1004	EGF-LIKE 25.
FT	DOMAIN	1006	1044	EGF-LIKE 26.
FT	DOMAIN	1046	1085	EGF-LIKE 27.
FT	DOMAIN	1087	1126	EGF-LIKE 28.
FT	DOMAIN	1130	1171	EGF-LIKE 29.
FT	DOMAIN	1472	1476	POLY-ARG.
FT	REPEAT	1165	1212	LIN/NOTCH 1.
FT	REPEAT	1213	1246	LIN/NOTCH 2.
FT	REPEAT	1247	1286	LIN/NOTCH 3.
FT	REPEAT	1633	1665	ANK 1.
FT	REPEAT	1666	1698	ANK 2.
FT	REPEAT	1700	1732	ANK 3.
FT	REPEAT	1733	1765	ANK 4.
FT	REPEAT	1766	1798	ANK 5.
FT	DISULFID	28	41	BY SIMILARITY.
FT	DISULFID	35	51	BY SIMILARITY.
FT	DISULFID	53	62	BY SIMILARITY.
FT	DISULFID	68	80	BY SIMILARITY.
FT	DISULFID	74	103	BY SIMILARITY.
FT	DISULFID	105	114	BY SIMILARITY.
FT	DISULFID	122	133	BY SIMILARITY.
FT	DISULFID	127	143	BY SIMILARITY.
FT	DISULFID	145	154	BY SIMILARITY.
FT	DISULFID	160	171	BY SIMILARITY.
FT	DISULFID	165	180	BY SIMILARITY.
FT	DISULFID	182	191	BY SIMILARITY.
FT	DISULFID	198	211	BY SIMILARITY.

Query Match 11.2%; Score 84.5; DB 1; Length 2003;
 Best Local Similarity 31.0%; Pred. No. 1.9;
 Matches 39; Conservative 6; Mismatches 42; Indels 39; Gaps 10;

QY	18	CAMLLCH-GS----	LOHTFQOHHLRP--	EGGTCEV----	IAAHRCCNKNRIEERSQTVK	66
Db	479	CLSQPCPGSTCLDLATP--	HCLCPGLEGLCEVETNECASAPCLNHADCHDLNGFQ	536		
QY	67	CSCLPKVGAGTTRNRPSC---	VDASIVIGKWCMEPCLEGE	-----	CKULPDNS	114
Db	537	CICLPG-FSGT-----	RCEEDIDE-----	CRSSPCANGGQCQDPQGFACKCLPGFE	582	
QY	115	GNWCAT	120			
Db	583	GPFCQT	588			

RESULT 15
 ID JAG2_MOUSE STANDARD; PRT; 1247 AA.
 AC Q9QY85; O55139; O70219;
 DT 28-FEB-2003 (Rel. 41, Created)
 DT 28-FEB-2003 (Rel. 41, Last sequence update)
 DT 15-MAR-2004 (Rel. 43, Last annotation update)
 DE Jagged 2 precursor (Jagged2).

GN JAG2.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 ON NCBI_TaxID=10090;
 OX [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Swiss Webster / NIH;
 RA Tsai S.;
 RL Submitted (DEC-1997) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP SEQUENCE OF 302-819 FROM N.A.
 RC TISSUE=Brain;
 RX MEDLINE=98051918; PubMed=9341252;
 RA Lan F., Jiang R., Shawber C., Weinmaster G., Gridley T.;
 RT "The Jagged2 gene maps to chromosome 12 and is a candidate for the lgl
 RT and sm mutations.";
 RL Mamm. Genome 8:875-876(1997).
 RN [3]
 RP SEQUENCE OF 325-759 FROM N.A.
 RC TISSUE=Brain;
 RX MEDLINE=98145947; PubMed=9486542;
 RA Valsecchi C., Ghezzi C., Ballabio A., Rugarli E.I.;
 RT "JAGGED2: a putative Notch ligand expressed in the apical ectodermal
 RT ridge and in sites of epithelial-mesenchymal interactions.";
 RL Mech. Dev. 69:203-207(1997).
 RN [4]
 RP TISSUE SPECIFICITY.
 RX MEDLINE=97459705; PubMed=9315665;
 RA Luo B., Aster J.C., Hasserjian R.P., Kuo F., Sklar J.;
 RT "Isolation and functional analysis of a cDNA for human Jagged2, a gene
 RT encoding a ligand for the Notch1 receptor.";
 RL Mol. Cell. Biol. 17:6057-6067(1997).
 CC -!- FUNCTION: Putative Notch ligand involved in the mediation of Notch
 CC signaling. Plays an essential role during limb, craniofacial and
 CC thymic development. May be involved in myogenesis and in the
 CC development of peripheral and central nervous systems.
 CC -!- SUBCELLULAR LOCATION: Type I membrane protein.
 CC -!- TISSUE SPECIFICITY: Found to be highest in fetal thymus,
 CC epidermis, foregut dorsal root ganglia and inner ear. In 2-week-
 CC old mice, abundant in heart, lung, thymus, skeletal muscle, brain
 CC and testis. Expression overlaps partially with Notch1 expression.
 CC -!- DEVELOPMENTAL STAGE: At 13 dpc, found in paravertebral vessels and
 CC dorsal root ganglia. At 14 dpc, in oropharyngeal epithelium,
 CC developing thymus and in the muscles of the tongue. By 15 dpc, in
 CC many tissues.
 CC -!- SIMILARITY: Contains 16 EGF-like domains.
 CC -!- SIMILARITY: Contains 1 DSL domain.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib-sib.ch).
 CC -----
 CC EMBL; AF038572; AAF16411.1; -;
 CC EMBL; AF010137; AAC14010.1; -;
 CC EMBL; Y14495; CAA74835.1; -;
 CC HSSP; P00743; ICCF.
 CC MGD; MGI:1098270; Jag2.
 CC GO; GO:0005887; C:integral to plasma membrane; IDA.
 CC GO; GO:0008083; F:growth factor activity; ISS.
 CC GO; GO:0005112; F:Notch binding; IPI.
 CC GO; GO:0007049; P:cell cycle; ISS.
 CC GO; GO:0030154; P:cell fate determination; NAS.
 CC GO; GO:0001709; P:cell fate determination; NAS.
 CC GO; GO:0007267; P:cell-cell signaling; NAS.
 CC GO; GO:0042492; P:gamma-delta T-cell differentiation; IMP.
 CC GO; GO:0009912; P:hair cell fate commitment; IMP.
 CC GO; GO:0007605; P:hearing; NAS.
 CC GO; GO:0030326; P:limb morphogenesis; ISS.
 CC -----

DR GO; GO:0007219; P:N signaling pathway; ISS.
 DR GO; GO:0030334; P:regulation of cell migration; ISS.
 DR GO; GO:0042127; P:regulation of cell proliferation; ISS.
 DR GO; GO:0007283; P:spermatogenesis; ISS.
 DR GO; GO:0045061; P:thymic T-cell selection; ISS.
 DR InterPro; IPR000152; Asx_hydroxyl_s.
 DR InterPro; IPR001774; DSL.
 DR InterPro; IPR000742; EGF_2.
 DR InterPro; IPR001881; EGF_Ca.
 DR InterPro; IPR001438; EGF_II.
 DR InterPro; IPR006209; EGF_like.
 DR InterPro; IPR001007; VWF_C.
 DR Pfam; PF01414; DSL; 1.
 DR Pfam; PF00008; EGF; 14.
 DR PRINTS; PR00010; EGFBL00D.
 DR SMART; SM00051; DSL; 1.
 DR SMART; SM00179; EGF_CA; 9.
 DR SMART; SM00214; VMC; 1.
 DR PROSITE; PS00010; ASX_HYDROXYL; 10.
 DR PROSITE; PS00022; EGF_1; 16.
 DR PROSITE; PS01186; EGF_2; 11.
 DR PROSITE; PS00026; EGF_3; 15.
 DR PROSITE; PS01187; EGF_CA; 7.
 KW Calcium-binding; EGF-like domain; Glycoprotein; Developmental protein;
 KW Repeat; Signal; Transmembrane.
 FT SIGNAL 1 23 POTENTIAL.
 FT CHAIN 24 1247 JAGGED 2.
 FT DOMAIN 24 1084 EXTRACELLULAR (POTENTIAL).
 FT TRANSMEM 1085 1107 POTENTIAL.
 FT DOMAIN 1108 1247 CYTOPLASMIC (POTENTIAL).
 FT DOMAIN 178 240 DSL.
 FT DOMAIN 241 274 EGF-LIKE 1.
 FT DOMAIN 275 305 EGF-LIKE 2.
 FT DOMAIN 307 345 EGF-LIKE 3.
 FT DOMAIN 347 383 EGF-LIKE 4.
 FT DOMAIN 385 421 EGF-LIKE 5.
 FT DOMAIN 423 459 EGF-LIKE 6.
 FT DOMAIN 461 496 EGF-LIKE 7.
 FT DOMAIN 498 534 EGF-LIKE 8.
 FT DOMAIN 536 572 EGF-LIKE 9.
 FT DOMAIN 574 634 EGF-LIKE 10 (ATYPICAL).
 FT DOMAIN 636 672 EGF-LIKE 11, CALCIUM-BINDING (POTENTIAL).
 FT DOMAIN 674 710 EGF-LIKE 12, CALCIUM-BINDING (POTENTIAL).
 FT DOMAIN 712 748 EGF-LIKE 13.
 FT DOMAIN 751 787 EGF-LIKE 14.
 FT DOMAIN 789 825 EGF-LIKE 15, CALCIUM-BINDING (POTENTIAL).
 FT DOMAIN 827 863 EGF-LIKE 16, CALCIUM-BINDING (POTENTIAL).
 FT DISULFID 245 256 BY SIMILARITY.
 FT DISULFID 249 262 BY SIMILARITY.
 FT DISULFID 264 273 BY SIMILARITY.
 FT DISULFID 276 287 BY SIMILARITY.
 FT DISULFID 282 293 BY SIMILARITY.
 FT DISULFID 295 304 BY SIMILARITY.
 FT DISULFID 311 323 BY SIMILARITY.
 FT DISULFID 317 333 BY SIMILARITY.
 FT DISULFID 335 344 BY SIMILARITY.
 FT DISULFID 351 362 BY SIMILARITY.
 FT DISULFID 356 371 BY SIMILARITY.
 FT DISULFID 373 382 BY SIMILARITY.
 FT DISULFID 389 400 BY SIMILARITY.
 FT DISULFID 394 409 BY SIMILARITY.
 FT DISULFID 411 420 BY SIMILARITY.
 FT DISULFID 427 438 BY SIMILARITY.
 FT DISULFID 432 447 BY SIMILARITY.
 FT DISULFID 449 458 BY SIMILARITY.
 FT DISULFID 465 475 BY SIMILARITY.
 FT DISULFID 469 484 BY SIMILARITY.
 FT DISULFID 486 495 BY SIMILARITY.
 FT DISULFID 502 513 BY SIMILARITY.
 FT DISULFID 507 522 BY SIMILARITY.
 FT DISULFID 524 533 BY SIMILARITY.
 FT DISULFID 540 551 BY SIMILARITY.
 FT DISULFID 545 560 BY SIMILARITY.

```

FT DISULFID 562 571 BY SIMILARITY.
FT DISULFID 589 612 POTENTIAL.
FT DISULFID 606 622 POTENTIAL.
FT DISULFID 624 633 BY SIMILARITY.
FT DISULFID 640 651 BY SIMILARITY.
FT DISULFID 645 660 BY SIMILARITY.
FT DISULFID 662 671 BY SIMILARITY.
FT DISULFID 678 689 BY SIMILARITY.
FT DISULFID 683 698 BY SIMILARITY.
FT DISULFID 700 709 BY SIMILARITY.
FT DISULFID 716 727 BY SIMILARITY.
FT DISULFID 721 736 BY SIMILARITY.
FT DISULFID 738 747 BY SIMILARITY.
FT DISULFID 755 766 BY SIMILARITY.
FT DISULFID 760 775 BY SIMILARITY.
FT DISULFID 777 786 BY SIMILARITY.
FT DISULFID 793 804 BY SIMILARITY.
FT DISULFID 798 813 BY SIMILARITY.
FT DISULFID 815 824 BY SIMILARITY.
FT DISULFID 831 842 BY SIMILARITY.
FT DISULFID 836 851 BY SIMILARITY.
FT DISULFID 853 862 BY SIMILARITY.
FT CARBOHYD 153 153 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 570 570 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 619 619 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 752 752 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 1060 1060 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CONFLICT 302 302 L -> M (IN REF. 2).
FT CONFLICT 461 461 N -> T (IN REF. 2).
FT CONFLICT 469 478 COHGTCKDL -> VSAWHLQGP (IN REF. 2).
FT CONFLICT 492 492 G -> V (IN REF. 2).
FT CONFLICT 546 546 L -> F (IN REF. 2).
FT CONFLICT 549 549 A -> V (IN REF. 2).
FT CONFLICT 735 738 RCAC -> PAR (IN REF. 3).
FT CONFLICT 809 809 N -> H (IN REF. 2).
FT CONFLICT 812 812 R -> A (IN REF. 2).
SQ SEQUENCE 1247 AA; 134726 MW; 1D80C8628FAFEEC CRC64;

Query Match 11.1%; Score 84; DB 1; Length 1247;
Best Local Similarity 24.6%; Pred. No. 1.3;
Matches 30; Conservative 12; Mismatches 30; Indels 50; Gaps 9;

QY 23 CHGSLQHTFQHHLRPEGGTC-EVIAAHC-----CNKNRI- 58
Db 465 CHGQCQH-----GGTCKDLVNGYCVCPRGFGRHCELEYDKCASSPCRRGGIC 513
QY 59 EERSQTVKCSCLPGKVAGTTRNRPS-CVDASIVICKWKCMEPCGECECKTLPDNGWM 117
Db 514 EDLVDGFRCHC-PRGLSGL-----HCEVDMDL-----CEPSPCLNGARCYNL--EGDY 559
QY 118 CA 119
Db 560 CA 561

```

Search completed: April 8, 2004, 13:16:40
Job time : 19 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 8, 2004, 13:15:04 ; Search time 22 Seconds
(without alignments)
312.102 Million cell updates/sec

Title: US-09-763-335-2
Perfect score: 754
Sequence: 1 MAMVSMSWVLYLWISACAM.....SGWMCATGNKIKTRIHPRT 133

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 389414 seqs, 51625971 residues

Total number of hits satisfying chosen parameters: 389414

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents AA:*
1: /cgn2_6/ptodata/2/iaa/5A_COMB.pep.*
2: /cgn2_6/ptodata/2/iaa/5B_COMB.pep.*
3: /cgn2_6/ptodata/2/iaa/6A_COMB.pep.*
4: /cgn2_6/ptodata/2/iaa/6B_COMB.pep.*
5: /cgn2_6/ptodata/2/iaa/ptotus_COMB.pep.*
6: /cgn2_6/ptodata/2/iaa/backfiles1.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	150	19.9	29	3	US-08-905-223-451
2	110.5	14.7	1248	3	US-08-892-046-6
3	96	12.7	1964	4	US-09-467-997-1
4	91.5	12.1	233	4	US-09-216-393B-110
5	86.5	11.5	156	4	US-09-252-991A-22898
6	86.5	11.5	1139	1	US-08-537-210A-4
7	86.5	11.5	1139	3	US-09-113-825-4
8	86.5	11.5	2703	1	US-08-185-432-19
9	86.5	11.5	2703	4	US-08-899-232-4
10	86	11.4	500	4	US-09-423-753-2
11	86	11.4	659	4	US-09-423-753-3
12	86	11.4	685	3	US-08-872-855-2
13	86	11.4	685	4	US-09-423-753-25
14	86	11.4	1010	3	US-08-882-046-7
15	86	11.4	1036	4	US-09-068-740A-6
16	86	11.4	1187	4	US-09-068-740A-7
17	86	11.4	1208	4	US-09-199-865-1
18	86	11.4	1218	2	US-08-400-159-6
19	86	11.4	1218	3	US-08-611-729A-6
20	86	11.4	1218	3	US-08-882-046-2
21	86	11.4	1218	3	US-09-214-278-7
22	86	11.4	1218	4	US-09-068-740A-11
23	86	11.4	1218	4	US-09-855-722-7
24	86	11.4	1219	3	US-08-882-046-5
25	84	11.1	2523	1	US-08-185-432-18
26	84	11.1	2523	4	US-08-899-232-3
27	82.5	10.9	2471	1	US-08-185-432-16

28	82.5	10.9	2471	1	US-08-083-590A-19
29	82.5	10.9	2471	3	US-08-532-384-19
30	82.5	10.9	2471	4	US-08-899-232-1
31	82	10.9	278	4	US-09-724-864-52
32	82	10.9	4544	1	US-08-469-486-52
33	82	10.9	4544	2	US-08-469-658-52
34	81.5	10.8	1148	3	US-08-882-046-4
35	81.5	10.8	2556	1	US-08-083-590A-20
36	81.5	10.8	2556	3	US-08-532-384-20
37	80.5	10.7	2321	4	US-09-230-652-2
38	80	10.6	2556	1	US-08-185-432-17
39	80	10.6	2556	4	US-08-899-232-2
40	79	10.5	1193	2	US-08-400-159-10
41	79	10.5	1193	3	US-08-611-729A-10
42	77.5	10.3	520	4	US-09-068-740A-3
43	77.5	10.3	702	4	US-09-068-740A-4
44	77.5	10.3	723	4	US-09-068-740A-9
45	77.5	10.3	723	4	US-09-423-753-27

ALIGNMENTS

RESULT 1
US-08-905-223-451
; Sequence 451, Application US/08905223
; Patent No. 6222029
; GENERAL INFORMATION:
; APPLICANT: Edwards, Jean-Baptiste D.
; APPLICANT: Duclert, Aymeric
; APPLICANT: Lacroix, Bruno
; TITLE OF INVENTION: 5' ESTS FOR SECRETED PROTEINS
; NUMBER OF SEQUENCES: 503
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Knobbe, Martens, Olson & Bear
; STREET: 501 West Broadway
; CITY: San Diego
; STATE: California
; COUNTRY: USA
; ZIP: 92101-3505
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy Disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: Win95
; SOFTWARE: Word
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/905,223
; FILING DATE:
; CLASSIFICATION: 536
; ATTORNEY/AGENT INFORMATION:
; NAME: Israel, Ned A.
; REGISTRATION NUMBER: 29,655
; REFERENCE/DOCKET NUMBER:
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (619) 235-8550
; TELEFAX: (619) 235-0176
; INFORMATION FOR SEQ ID NO: 451:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 29 amino acids
; TYPE: AMINO ACID
; TOPOLOGY: LINEAR
; MOLECULE TYPE: PROTEIN
; ORIGINAL SOURCE:
; ORGANISM: Homo Sapiens
; TISSUE TYPE: Brain
; FEATURE:
; NAME/KEY: sig_peptide
; LOCATION: -25...-1
; IDENTIFICATION METHOD: Von Heijne matrix
; OTHER INFORMATION: score 5.8
; OTHER INFORMATION: seq LWSACAMLLCHG/SL
US-08-905-223-451

```

Query Match      19.9%; Score 150; DB 3; Length 29;
Best Local Similarity 100.0%; Pred. No. 1.9e-09;
Matches 28; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MAMVSAMSWLYLWISACAMLLCHGSLQ 28
Db 1 MAMVSAMSWLYLWISACAMLLCHGSLQ 28

RESULT 2
US-08-882-046-6
; Sequence 6, Application US/08882046
; Patent No. 6136952
; GENERAL INFORMATION:
; APPLICANT: Li, Linheng
; APPLICANT: Hood, Leroy
; APPLICANT: Krantz, Ian D.
; APPLICANT: Spinner, Nancy B.
; TITLE OF INVENTION: Human Jagged Polypeptide, Encoding
; TITLE OF INVENTION: Nucleic Acids and Methods of Use
; NUMBER OF SEQUENCES: 110
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Campbell & Flores LLP
; STREET: 4370 La Jolla Village Drive, Suite 700
; CITY: San Diego
; STATE: California
; COUNTRY: USA
; ZIP: 92122
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/882,046
; FILING DATE: 25-JUN-1997
; CLASSIFICATION: 536
; ATTORNEY/AGENT INFORMATION:
; NAME: Campbell, Cathryn A.
; REGISTRATION NUMBER: 31,815
; REFERENCE/DOCKET NUMBER: P-UW 2637
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (619) 535-9001
; TELEFAX: (619) 535-8949
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1248 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-08-882-046-6

Query Match      14.7%; Score 110.5; DB 3; Length 1248;
Best Local Similarity 27.0%; Pred. No. 0.0022;
Matches 33; Conservative 13; Mismatches 27; Indels 49; Gaps 9;

QY 23 CHGSLQHTFQOHHLRPEGGTC-EVIAAHR-----CNKNRI- 58
Db 465 CHGQCQH-----GGTKDLVNGYQVCPRGFGGRHCELEYKCASSPCRRGGIC 513

QY 59 EERSQTVKCSCLPGKVGAGTTRNRPS-CVDASIVIGKWWCEMEPCLEGECKTLPDNSGWM 117
Db 514 EBLVDGFRCHC-FRGLSG-----PLCEVDVDL-----WCEPNPCLNGARCYNLEDD---YY 560

QY 118 CA 119
Db 561 CA 562

RESULT 3
US-09-467-997-1
; Sequence 1, Application US/09467997
; Patent No. 6379925

```

```

; GENERAL INFORMATION:
; APPLICANT: Kitajewski, Jan
; APPLICANT: Uyttendaele, Hendrik
; TITLE OF INVENTION: ANGIOGENIC MODULATION BY NOTCH SIGNAL TRANSDUCTION
; FILE REFERENCE: 53863-A-PCT-US
; CURRENT APPLICATION NUMBER: US/09/467,997
; CURRENT FILING DATE: 1999-12-20
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
; LENGTH: 1964
; TYPE: PRT
; ORGANISM: mouse
US-09-467-997-1

```

```

Query Match      12.7%; Score 96; DB 4; Length 1964;
Best Local Similarity 26.8%; Pred. No. 0.13;
Matches 33; Conservative 8; Mismatches 48; Indels 34; Gaps 6;

QY 15 ISACAMLLC-HGSLQHTFQOHHLRPEGGTC-----EVIAAHR--CCNKNRI 58
Db 689 LGGCISTPCAAGGTGCHP-----QPSGYNCTCPAGYMGILCTSEEVTAACHSGPCLNGGSC 741

QY 59 EERSQTVKCSCLPGKVGAGTTRNRPS-CVDASIVIGKWWCEMEPCLEGECKTLPDNSGWM 118
Db 742 SIRPEGYSTCLP---SHTGRHCQTAVD-----HCVSASCLNGGTCVNKEGTFFCLC 790

QY 119 ATG 121
Db 791 ATG 793

```

```

RESULT 4
US-09-216-393B-110
; Sequence 110, Application US/09216393B
; Patent No. 6514694
; GENERAL INFORMATION:
; APPLICANT: Milhausen, Michael James
; TITLE OF INVENTION: TOXOPLASMA GONDII PROTEINS, NUCLEIC ACID MOLECULES, AND USES THEREOF
; FILE REFERENCE: TX-1-C2
; CURRENT APPLICATION NUMBER: US/09/216,393B
; CURRENT FILING DATE: 1998-12-18
; PRIOR APPLICATION NUMBER: 08/994,825
; PRIOR FILING DATE: 1997-12-19
; NUMBER OF SEQ ID NOS: 366
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 110
; LENGTH: 233
; TYPE: PRT
; ORGANISM: Toxoplasma gondii
US-09-216-393B-110

```

```

Query Match      12.1%; Score 91.5; DB 4; Length 233;
Best Local Similarity 25.9%; Pred. No. 0.036;
Matches 28; Conservative 6; Mismatches 39; Indels 35; Gaps 4;

QY 18 CAMLLCHGSLQHTFQOHHLRPEGGTC-EVIAAHR-----CNKN-----RIERS 62
Db 78 CQDCACHG-----GKTCETTKHCCLNDSCNGHGTCTNTNTNCCEAGF 122

QY 63 QTVKCSCLPGKVGAGTTRNRPS-CVDASIVIGKWWCEMEPCLEGECKTL 110
Db 123 AGTNCSSSEKCSGKT-----CLSGHCNPATGACVCDPCHTGERCETL 165

```

```

RESULT 5
US-09-252-991A-22898
; Sequence 22898, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS

```



```
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; PRIOR FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 22898
; LENGTH: 156
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-22898

Query Match      11.5%; Score 86.5; DB 4; Length 156;
Best Local Similarity 25.0%; Pred. No. 0.078;
Matches 32; Conservative 9; Mismatches 64; Indels 23; Gaps 5;

QY 6 AMSWLVLYMISACAMLLCHGSLQHTQOHLHREPGTCEVIAAHRCCNKNRIEERSQTV 65
Db 12 AASW--KSWSSACS--TSTGCLPTYAPARRPSPARGSSSKAVAAARCCNATTPCSNSRSP 67
QY 66 K-----GSCLPKGVAGTTR-----NRPSQVDASIVGKWCMEPCLEGECKT 109
Db 68 NPCCLCWSSGTCRCCLPISTDRMTPTASATRESTPSAPV---RWRPRPACTSMRCSR 124
QY 110 LPNSGWM 117
Db 125 RSARRGWI 132

RESULT 6
US-08-537-210A-4
; Sequence 4, Application US/08537210A
; Patent No. 5780300
; GENERAL INFORMATION:
; APPLICANT: Artavanis-Tsakonas, Spyridon
; APPLICANT: Fortini, Mark
; APPLICANT: Matsuno, Kenji
; TITLE OF INVENTION: MANIPULATION OF NON-TERMINALLY
; TITLE OF INVENTION: DIFFERENTIATED CELLS USING THE NOTCH PATHWAY
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Pennie & Edmonds
; STREET: 1155 Avenue of the Americas
; CITY: New York
; STATE: NY
; COUNTRY: USA
; ZIP: 10036/2711
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/537,210A
; FILING DATE: 29-SEP-1995
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Misrock, S. Leslie
; REGISTRATION NUMBER: 18,872
; REFERENCE/DOCKET NUMBER: 7326-027
; TELEPHONE: 212-790-9090
; TELEFAX: 212-869-8864
; TELEX: 66141 PENNIE
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1139 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: unknown
; MOLECULE TYPE: protein
```

```
; FEATURE:
; NAME/KEY: Dros N
; LOCATION: 1189...2327
; OTHER INFORMATION: Highly conserved ankyrin repeat
; OTHER INFORMATION: region of No. 5780300ch
US-08-537-210A-4

Query Match      11.5%; Score 86.5; DB 1; Length 1139;
Best Local Similarity 25.9%; Pred. No. 0.74;
Matches 22; Conservative 10; Mismatches 34; Indels 19; Gaps 3;

QY 40 EGGTCEV-----IAAHRCCNKNRIEERSQTVKSCCLPGKVA-----GTTNRPS 83
Db 26 QQNCNELNIDDCAPNCPGGTCHDRVMNFSCPPGTGWTGIIICEINKDKCKPGACHNNGS 85
QY 84 CYDASIVGKWCMEPCLEGECK 108
Db 86 CIDR---VGFECVCPQGFVGARCE 107

RESULT 7
US-09-113-825-4
; Sequence 4, Application US/09113825
; Patent No. 6149902
; GENERAL INFORMATION:
; APPLICANT: Artavanis-Tsakonas, Spyridon
; APPLICANT: Fortini, Mark
; APPLICANT: Matsuno, Kenji
; TITLE OF INVENTION: MANIPULATION OF NON-TERMINALLY
; TITLE OF INVENTION: DIFFERENTIATED CELLS USING THE NOTCH PATHWAY
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Pennie & Edmonds
; STREET: 1155 Avenue of the Americas
; CITY: New York
; STATE: NY
; COUNTRY: USA
; ZIP: 10036/2711
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/113,825
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/537,210
; FILING DATE: 29-SEP-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Misrock, S. Leslie
; REGISTRATION NUMBER: 18,872
; REFERENCE/DOCKET NUMBER: 7326-027
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212-790-9090
; TELEFAX: 212-869-8864
; TELEX: 66141 PENNIE
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1139 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: unknown
; MOLECULE TYPE: protein
; FEATURE:
; NAME/KEY: Dros N
; LOCATION: 1189...2327
; OTHER INFORMATION: Highly conserved ankyrin repeat
; OTHER INFORMATION: region of No. 6149902ch
US-09-113-825-4

Query Match      11.5%; Score 86.5; DB 3; Length 1139;
```

```
Best Local Similarity 25.9%; Pred. No. 0.74;
Matches 22; Conservative 10; Mismatches 34; Indels 19; Gaps 3;

QY 40 EGGTCV-----IAAHRCCNKNRIEERSQTVKCSCLPGKVA-----GTTNRPS 83
Db 26 QQONCELNIDDCAPNCPQGGTCHDRVMNFSCPPGTGIIICEINKDKCKPGACHNNGS 85
QY 84 CVDASIVIGKWCMEPCLEGECK 108
Db 86 CIDR---VGGFECVCQGFVGARCE 107

RESULT 8
US-08-185-432-19
; Sequence 19, Application US/08185432
; Patent No. 5750652
; GENERAL INFORMATION:
; APPLICANT: Artavanis-Tsakonas, Spyridon
; APPLICANT: Busseau, Isabelle
; APPLICANT: Diederich, Robert J.
; APPLICANT: Xu, Tian
; APPLICANT: Matsuno, Kenji
; TITLE OF INVENTION: DELTEX PROTEINS, NUCLEIC ACIDS, AND
; TITLE OF INVENTION: ANTIBODIES, AND RELATED METHODS AND COMPOSITIONS
; NUMBER OF SEQUENCES: 23
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: PENNIE & EDMONDS
; STREET: 1155 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: U.S.A.
; ZIP: 10036-2711
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA: US/08/185,432
; FILING DATE: 21-JAN-1994
; CLASSIFICATION: 530
; ATTORNEY/AGENT INFORMATION:
; NAME: Misrock, S. Leslie
; REGISTRATION NUMBER: 18,872
; REFERENCE/DOCKET NUMBER: 7326-006
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 790-9090
; TELEFAX: (212) 869-8864/9741
; TELEX: 66141 PENNIE
; INFORMATION FOR SEQ ID NO: 19:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2703 amino acids
; TYPE: amino acid
; TOPOLOGY: unknown
; MOLECULE TYPE: protein
US-08-185-432-19

Query Match 11.5%; Score 86.5; DB 1; Length 2703;
Best Local Similarity 25.9%; Pred. No. 2;
Matches 22; Conservative 10; Mismatches 34; Indels 19; Gaps 3;

QY 40 EGGTCV-----IAAHRCCNKNRIEERSQTVKCSCLPGKVA-----GTTNRPS 83
Db 1214 QQONCELNIDDCAPNCPQGGTCHDRVMNFSCPPGTGIIICEINKDKCKPGACHNNGS 1273
QY 84 CVDASIVIGKWCMEPCLEGECK 108
Db 1274 CIDR---VGGFECVCQGFVGARCE 1295

RESULT 9
US-08-899-232-4
; Sequence 4, Application US/08899232
```

```
; Patent No. 643650
; GENERAL INFORMATION:
; APPLICANT: Artavanis-Tsakonas, Spyridon
; APPLICANT: Qi, Huilin
; TITLE OF INVENTION: ACTIVATED FORMS OF NOTCH AND METHODS BASED THEREON
; FILE REFERENCE: 7326-046
; CURRENT APPLICATION NUMBER: US/08/899,232
; CURRENT FILING DATE: 1997-07-23
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 4
; LENGTH: 2703
; TYPE: PRT
; ORGANISM: Drosophila sp.
US-08-899-232-4

Query Match 11.5%; Score 86.5; DB 4; Length 2703;
Best Local Similarity 25.9%; Pred. No. 2;
Matches 22; Conservative 10; Mismatches 34; Indels 19; Gaps 3;

QY 40 EGGTCV-----IAAHRCCNKNRIEERSQTVKCSCLPGKVA-----GTTNRPS 83
Db 1214 QQONCELNIDDCAPNCPQGGTCHDRVMNFSCPPGTGIIICEINKDKCKPGACHNNGS 1273
QY 84 CVDASIVIGKWCMEPCLEGECK 108
Db 1274 CIDR---VGGFECVCQGFVGARCE 1295

RESULT 10
US-09-423-753-2
; Sequence 2, Application US/09423753
; Patent No. 6664098
; GENERAL INFORMATION:
; APPLICANT: SAKANO, SEIJI
; TITLE OF INVENTION: NOVEL DIFFERENTIATION INHIBITOR
; FILE REFERENCE: KP-8693
; CURRENT APPLICATION NUMBER: US/09/423,753
; CURRENT FILING DATE: 1999-12-30
; PRIOR APPLICATION NUMBER: PCT/JP98/02104
; PRIOR FILING DATE: 1998-05-13
; PRIOR APPLICATION NUMBER: JP 9/124064
; PRIOR FILING DATE: 1997-05-14
; NUMBER OF SEQ ID NOS: 27
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
; LENGTH: 500
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-423-753-2

Query Match 11.4%; Score 86; DB 4; Length 500;
Best Local Similarity 25.0%; Pred. No. 0.33;
Matches 28; Conservative 9; Mismatches 45; Indels 30; Gaps 6;

QY 14 WISAC-----AMLLCHGSLQHTFQHHLHREPGTCEVIAAHRCCNKNRIEERSQTVKCS 69
Db 242 WQCTDEGHWGLFCDQDLNYC---THSPCKNGATC-----SNSGQSYT---CTC 286
QY 70 LPGKVAGTTRNRPSCVDASIVIGKWCMEPCLEGECKTLPDNSGWMCATG 121
Db 287 RPYVTG-----VDCLELSE--CDSNPCRNCGSKQDEGYHCLCPPG 327

RESULT 11
US-09-423-753-3
; Sequence 3, Application US/09423753
; Patent No. 6664098
; GENERAL INFORMATION:
; APPLICANT: SAKANO, SEIJI
; TITLE OF INVENTION: NOVEL DIFFERENTIATION INHIBITOR
; FILE REFERENCE: KP-8693
; CURRENT APPLICATION NUMBER: US/09/423,753
```

; CURRENT FILING DATE: 1999-12-30
; PRIOR APPLICATION NUMBER: PCT/JP98/02104
; PRIOR FILING DATE: 1998-05-13
; PRIOR APPLICATION NUMBER: JP 9/124064
; PRIOR FILING DATE: 1997-05-14
; NUMBER OF SEQ ID NOS: 27
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 3
; LENGTH: 659
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-423-753-3

Query Match 11.4%; Score 86; DB 4; Length 659;
Best Local Similarity 25.0%; Pred. No. 0.45;
Matches 28; Conservative 9; Mismatches 45; Indels 30; Gaps 6;

QY 14 WISAC-----AMLLCHGSLQHTFOQHHLRPEGGTCEVIAAHRCCNKNRIERSQTVKCS 69
Db 242 WQCTDEGWLFCDDQDLYNC--THSPCKNGATC-----SNSGQSYT--CTC 286

QY 70 LPGAAGTTNRPCVNDASIVIGKWCMEPCLEGECKTLPDNGMWCATG 121
Db 287 RPYGTG-----VDCELELSE--CDNPNCRNGSGCKQEDGYHCLCPPG 327

RESULT 12
US-08-872-855-2
; Sequence 2, Application US/08872855
; Patent No. 6121045
; GENERAL INFORMATION:
; APPLICANT: McCarthy, Sean
; APPLICANT: Gearing, David
; TITLE OF INVENTION: NOVEL HUMAN DELTA3 COMPOSITIONS AND
; TITLE OF INVENTION: THERAPEUTIC USES THEREFOR
; NUMBER OF SEQUENCES: 23
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: FOLEY, HOAG & ELIOT LLP
; STREET: One Post Office Square
; CITY: Boston
; STATE: MA
; COUNTRY: USA
; ZIP: 02109-2170
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION NUMBER: US/08/872,855
; FILING DATE: 11-JUN-1997
; CLASSIFICATION: 514
; ATTORNEY/AGENT INFORMATION:
; NAME: Arnold, Beth E.
; REGISTRATION NUMBER: 35,430
; REFERENCE/DOCKET NUMBER: MAA-003.02
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617-832-1000
; TELEFAX: 617-832-7000
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 685 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-872-855-2

Query Match 11.4%; Score 86; DB 3; Length 685;
Best Local Similarity 25.0%; Pred. No. 0.47;
Matches 28; Conservative 9; Mismatches 45; Indels 30; Gaps 6;

QY 14 WISAC-----AMLLCHGSLQHTFOQHHLRPEGGTCEVIAAHRCCNKNRIERSQTVKCS 69

Db 268 WQCTDEGWLFCDDQDLYNC--THSPCKNGATC-----SNSGQSYT--CTC 312

QY 70 LPGAAGTTNRPCVNDASIVIGKWCMEPCLEGECKTLPDNGMWCATG 121
Db 313 RPYGTG-----VDCELELSE--CDNPNCRNGSGCKQEDGYHCLCPPG 353

RESULT 13
US-09-423-753-25
; Sequence 25, Application US/09423753
; Patent No. 6664098
; GENERAL INFORMATION:
; APPLICANT: SAKANO, SEIJI
; TITLE OF INVENTION: NOVEL DIFFERENTIATION INHIBITOR
; FILE REFERENCE: KP-8693
; CURRENT APPLICATION NUMBER: US/09/423,753
; CURRENT FILING DATE: 1999-12-30
; PRIOR APPLICATION NUMBER: PCT/JP98/02104
; PRIOR FILING DATE: 1998-05-13
; PRIOR APPLICATION NUMBER: JP 9/124064
; PRIOR FILING DATE: 1997-05-14
; NUMBER OF SEQ ID NOS: 27
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 25
; LENGTH: 685
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-423-753-25

Query Match 11.4%; Score 86; DB 4; Length 685;
Best Local Similarity 25.0%; Pred. No. 0.47;
Matches 28; Conservative 9; Mismatches 45; Indels 30; Gaps 6;

QY 14 WISAC-----AMLLCHGSLQHTFOQHHLRPEGGTCEVIAAHRCCNKNRIERSQTVKCS 69
Db 268 WQCTDEGWLFCDDQDLYNC--THSPCKNGATC-----SNSGQSYT--CTC 312

QY 70 LPGAAGTTNRPCVNDASIVIGKWCMEPCLEGECKTLPDNGMWCATG 121
Db 313 RPYGTG-----VDCELELSE--CDNPNCRNGSGCKQEDGYHCLCPPG 353

RESULT 14
US-08-882-046-7
; Sequence 7, Application US/08882046
; Patent No. 6136952
; GENERAL INFORMATION:
; APPLICANT: Li, Linheng
; APPLICANT: Hood, Leroy
; APPLICANT: Krantz, Ian D.
; APPLICANT: Spinner, Nancy B.
; TITLE OF INVENTION: Human Jagged Polypeptide, Encoding
; TITLE OF INVENTION: Nucleic Acids and Methods of Use
; NUMBER OF SEQUENCES: 110
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Campbell & Flores LLP
; STREET: 4370 La Jolla Village Drive, Suite 700
; CITY: San Diego
; STATE: California
; COUNTRY: USA
; ZIP: 92122
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION NUMBER: US/08/882,046
; FILING DATE: 25-JUN-1997
; CLASSIFICATION: 536
; ATTORNEY/AGENT INFORMATION:
; NAME: Campbell, Cathryn A.
; REGISTRATION NUMBER: 31,815

```

: REFERENCE/DOCKET NUMBER:  P-UW 2637
:
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: (619) 535-3001
: TELEFAX: (619) 535-8949
: INFORMATION FOR SEQ ID NO: 7:
: SEQUENCE CHARACTERISTICS
: LENGTH: 1010 amino acids
: TYPE: amino acid
: TOPOLOGY: linear
: MOLECULE TYPE: protein
: US-08-882-046-7

```

```

Query Match      11.4%; Score 86; DB 3; Length 1010;
Best Local Similarity 23.8%; Pred.No. 0.73;
Matches 34; Conservative 11; Mismatches 34; Indels 64; Gaps 9;

QY    37 HRP-EGTCE-----VIAAHC-----C-NKNRIERSQTVMKSC 69
       ||| |||| |
Db     303 HQPCLNGGTCSNTGPDKYQCSPGYSGPNCEIAEHACLSDPCHNRSGSKETSLGFEEC 362
               :|: ||| |
QY    70 LPGKVAGTRNR-----PSCVDASIVIG-----CEMEP 100
               :|: ||| |
Db     363 SPGWGTGPTCSNIDDCSPNNCSHGTCQD--LVANGFKVCVPQWTKTQLCDANECEAKP 420
               :|: ||| |
QY    101 CLEGECKTLFDN-----SGWM 117
               :|: ||| |
Db     421 CVNAKSCKNIASIYYCDLCPGMW 443

```

```

RESULT 15
US-09-068-740A-6
; Sequence 6, Application US/09068740A
; Patent No. 6337387
; GENERAL INFORMATION:
; APPLICANT: SAKANO, SEIJI
; APPLICANT: ITOH, AKIRA
; TITLE OF INVENTION: DIFFERENTIATION-SUPPRESSIVE POLYPEPTIDE
; FILE REFERENCE: KP-8447
; CURRENT APPLICATION NUMBER: US/09/068,740A
; CURRENT FILING DATE: 1998-06-18
; PRIOR APPLICATION NUMBER: JP 7-299611
; PRIOR FILING DATE: 1995-11-17
; PRIOR APPLICATION NUMBER: JP 7-311811
; PRIOR FILING DATE: 1995-11-30
; PRIOR APPLICATION NUMBER: PCT/JF96/03356
; PRIOR FILING DATE: 1996-11-15
; NUMBER OF SEQ ID NOS: 48
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 6
; LENGTH: 1036
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-068-740A-6

```

```

Query Match      11.4%; Score 86; DB 4; Length 1036;
Best Local Similarity 23.8%; Pred. No. 0.75;
Matches 34; Conservative 11; Mismatches 34; Indels 64; Gaps 9;

QY      37 HRP--EGTCE-----VIAAHR-----C-NKRIERSOTVKSC 69
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db      272 HQCLNGTGSNTGFDKYQCSPEYSGPNCEIAEHACLSDPCHNRGSKETSLGFCEC 331
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

QY      70 LPQKVAAGTTRNR-----PSCVDASIVIG-----KWW-----CEMEP 100
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db      332 SPGWGTGFTCTNIDDCSPNNCSHGTCQD--LVNGFKVCVCPQWTKTCLQDANECEAKP 389
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

QY      101 CLGEECKTLPDN-----SGWM 117
      ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db      390 CVNAKSKNLIASYYCDLPGWM 412
      ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

```

Search completed: April 8, 2004, 13:18:40
Job time : 23 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 8, 2004, 13:14:59 ; Search time 20 Seconds
(without alignments)
639.674 Million cell updates/sec

Title: US-09-763-335-2
Perfect score: 754
Sequence: 1 MAVVSMSWVLYLWISACAM.....SGMWCATGNKIKYTRIHPRT 133

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283366 seqs, 96191526 residues

Total number of hits satisfying chosen parameters: 283366

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Database :

PIR 78:*
1: Pirl:*
2: Pirl:*
3: Pirl:*
4: Pirl:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	106	14.1	3871	2 T22812	hypothetical prote
2	96	12.7	1964	2 T09059	notch4 - mouse
3	95.5	12.7	2531	2 S18188	notch protein homo
4	95.5	12.7	2531	2 A46019	notch-1 protein -
5	94.5	12.5	2352	2 T30201	Notch homolog prot
6	93	12.3	2531	2 T31070	notch homolog - se
7	92.5	12.3	2139	2 A35672	crumbs protein - f
8	86.5	11.5	2703	1 A24420	notch protein - fr
9	86	11.4	685	2 JC7570	Delta-4 protein -
10	86	11.4	1220	2 A56136	jagged protein pre
11	85.5	11.3	1064	2 A40136	fibropellin Ia - s
12	84	11.1	2524	2 A35844	Xotch protein - Af
13	82.5	10.9	1203	2 A49175	Notch B protein -
14	82.5	10.9	2471	2 A49128	cell-fate determin
15	82	10.9	2318	2 S45306	notch 3 protein -
16	82	10.9	4544	1 S02392	alpha-2-macroglobu
17	82	10.9	5376	2 T42215	zonadhesin - mouse
18	81.5	10.8	355	2 S56058	probable membrane
19	81.5	10.8	2555	2 A40043	notch protein homo
20	81	10.7	907	2 T27317	hypothetical prote
21	80.5	10.7	2321	2 S78549	notch3 protein - h
22	79	10.5	293	2 B26637	neurogenic repetit
23	79	10.5	655	2 JC7850	disintegrin and me
24	79	10.5	660	2 S71949	metalloproteinase
25	79	10.5	861	2 A48825	Notch homolog Motc
26	78.5	10.4	502	2 T20130	hypothetical prote
27	78.5	10.4	838	2 T20125	hypothetical prote
28	78.5	10.4	1178	1 A39804	thrombospondin pre
29	78.5	10.4	1364	2 T00250	MSGF2 protein - hu

30	78	10.3	2918	2 A54105	fibrillin-2 precur
31	78	10.3	4543	1 A53102	alpha-2-macroglobu
32	77.5	10.3	686	2 JC7569	Delta-4 protein -
33	77.5	10.3	1125	1 S57846	protein-tyrosine k
34	77.5	10.3	1548	2 S34583	serine proteinase
35	77.5	10.3	2437	2 S42612	transmembrane prot
36	77	10.2	843	1 A27340	complement C7 prec
37	77	10.2	1192	2 S69000	laminin gamma 2 ch
38	77	10.2	2907	2 A57278	fibrillin-2 precur
39	77	10.2	2946	2 T15840	hypothetical prote
40	76.5	10.1	1172	2 A42587	thrombospondin 2 p
41	76	10.1	941	1 A55195	chordin precursor
42	75.5	10.0	1722	2 B89753	protein FilC7.4 li
43	75.5	10.0	2871	2 A55624	fibrillin-1 precur
44	75	9.9	473	2 A56175	adhesive plaque pr
45	75	9.9	728	2 I50719	C-Delta-1 - chicke

ALIGNMENTS

RESULT 1

T22812
hypothetical protein ZC116.3 - Caenorhabditis elegans

C;Species: Caenorhabditis elegans
C;Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 16-Aug-2002
C;Accession: T22812; T27494

R;Burton, J.
submitted to the EMBL Data Library, June 1996

A;Reference number: Z19618
A;Accession: T22812

A;Status: preliminary; translated from GB/EMBL/DBJ
A;Molecule type: DNA

A;Residues: 1-3871 <WIL>
A;Cross-references: EMBL:Z74473; PIDN:CRAA98952.1; GSPDB:GN00023; CESP:ZC116.3

A;Experimental source: clone F56H9
R;Smye, R.
submitted to the EMBL Data Library, June 1996

A;Reference number: Z20376
A;Accession: T27494

A;Status: preliminary; translated from GB/EMBL/DBJ
A;Molecule type: DNA

A;Residues: 1-3871 <WIL>
A;Cross-references: EMBL:Z74046; PIDN:CRAA98557.1; GSPDB:GN00023; CESP:ZC116.3

A;Experimental source: clone ZC116
C;Genetics:
A;Gene: CESP:ZC116.3

A;Map position: 5
A;Introns: 29/3; 66/2; 244/1; 359/3; 422/1; 549/1; 572/1; 728/1; 820/1; 889/3; 1189/1; 12763/1; 2809/1; 2906/3; 3051/3; 3176/1; 3254/1; 3341/2; 3536/1; 3585/2; 3582/3; 3609/1; 3609/1;

C;Superfamily: intrinsic factor-B12 receptor cubilin; EGF homology
Query Match 14.1%; Score 106; DB 2; Length 3871;
Best Local Similarity 31.2%; Pred. No. 0.069;
Matches 34; Conservative 13; Mismatches 40; Indels 22; Gaps 6;

Qy 24 HGSLSHTFQHHUHPGGTCVIAHRCNKNRIERSQ-----TVKCSCLPGKVAGT 77
Db 148 HGSLS-----CQYHMSAC-SKTFELCGPHGHCIESIVDPTGSSSDTTTTCICDWGFKVSS 202

Qy 78 TRNRPSVDASIVIGKWCMEPCLEGECKTLPDN---SGMWCATGNK 123
Db 203 DRNNPTCDVNV-----ECESNFCHPGVDICINLPGSFVCSG--CPKGYK 243

RESULT 2

T09059
notch4 - mouse

C;Species: Mus musculus (house mouse)
C;Date: 11-Jun-1999 #sequence_revision 11-Jun-1999 #text_change 08-Sep-2002
C;Accession: T09059

R;Rowen, L.; Mahairas, G.; Qin, S.; Ahearn, M.E.; Dankers, C.; Lasky, C.; Sc
submitted to the EMBL Data Library, October 1997

A;Description: Sequence of the mouse major histocompatibility locus class III region.
A;Reference number: Z16543
A;Accession: T09059
A;Status: preliminary; translated from GB/EMBL/DBJ
A;Molecule type: DNA
A;Residues: 1-1964 <ROW>
A;Cross-references: EMBL:AF030001; NID:g2564945; PID:g2564947
C;Genetics:
A;Gene: notch4
A;Map position: 17
A;Introns: 22/1; 49/2; 148/1; 264/1; 305/1; 384/1; 436/1; 501/1; 539/1; 577/1; 618/1; 676/1; 724/1; 772/1; 811/1; 841/1; 871/1; 901/1; 931/1; 961/1; 991/1; 1021/1; 1051/1; 1081/1; 1111/1; 1141/1; 1171/1; 1201/1; 1231/1; 1261/1; 1291/1; 1321/1; 1351/1; 1381/1; 1411/1; 1441/1; 1471/1; 1501/1; 1531/1; 1561/1; 1591/1; 1621/1; 1651/1; 1681/1; 1711/1; 1741/1; 1771/1; 1801/1; 1831/1; 1861/1; 1891/1; 1921/1; 1951/1; 1981/1; 2011/1; 2041/1; 2071/1; 2101/1; 2131/1; 2161/1; 2191/1; 2221/1; 2251/1; 2281/1; 2311/1; 2341/1; 2371/1; 2401/1; 2431/1; 2461/1; 2491/1; 2521/1; 2551/1; 2581/1; 2611/1; 2641/1; 2671/1; 2701/1; 2731/1; 2761/1; 2791/1; 2821/1; 2851/1; 2881/1; 2911/1; 2941/1; 2971/1; 3001/1; 3031/1; 3061/1; 3091/1; 3121/1; 3151/1; 3181/1; 3211/1; 3241/1; 3271/1; 3301/1; 3331/1; 3361/1; 3391/1; 3421/1; 3451/1; 3481/1; 3511/1; 3541/1; 3571/1; 3601/1; 3631/1; 3661/1; 3691/1; 3721/1; 3751/1; 3781/1; 3811/1; 3841/1; 3871/1; 3901/1; 3931/1; 3961/1; 3991/1; 4021/1; 4051/1; 4081/1; 4111/1; 4141/1; 4171/1; 4201/1; 4231/1; 4261/1; 4291/1; 4321/1; 4351/1; 4381/1; 4411/1; 4441/1; 4471/1; 4501/1; 4531/1; 4561/1; 4591/1; 4621/1; 4651/1; 4681/1; 4711/1; 4741/1; 4771/1; 4801/1; 4831/1; 4861/1; 4891/1; 4921/1; 4951/1; 4981/1; 5011/1; 5041/1; 5071/1; 5101/1; 5131/1; 5161/1; 5191/1; 5221/1; 5251/1; 5281/1; 5311/1; 5341/1; 5371/1; 5401/1; 5431/1; 5461/1; 5491/1; 5521/1; 5551/1; 5581/1; 5611/1; 5641/1; 5671/1; 5701/1; 5731/1; 5761/1; 5791/1; 5821/1; 5851/1; 5881/1; 5911/1; 5941/1; 5971/1; 6001/1; 6031/1; 6061/1; 6091/1; 6121/1; 6151/1; 6181/1; 6211/1; 6241/1; 6271/1; 6301/1; 6331/1; 6361/1; 6391/1; 6421/1; 6451/1; 6481/1; 6511/1; 6541/1; 6571/1; 6601/1; 6631/1; 6661/1; 6691/1; 6721/1; 6751/1; 6781/1; 6811/1; 6841/1; 6871/1; 6901/1; 6931/1; 6961/1; 6991/1; 7021/1; 7051/1; 7081/1; 7111/1; 7141/1; 7171/1; 7201/1; 7231/1; 7261/1; 7291/1; 7321/1; 7351/1; 7381/1; 7411/1; 7441/1; 7471/1; 7501/1; 7531/1; 7561/1; 7591/1; 7621/1; 7651/1; 7681/1; 7711/1; 7741/1; 7771/1; 7801/1; 7831/1; 7861/1; 7891/1; 7921/1; 7951/1; 7981/1; 8011/1; 8041/1; 8071/1; 8101/1; 8131/1; 8161/1; 8191/1; 8221/1; 8251/1; 8281/1; 8311/1; 8341/1; 8371/1; 8401/1; 8431/1; 8461/1; 8491/1; 8521/1; 8551/1; 8581/1; 8611/1; 8641/1; 8671/1; 8701/1; 8731/1; 8761/1; 8791/1; 8821/1; 8851/1; 8881/1; 8911/1; 8941/1; 8971/1; 9001/1; 9031/1; 9061/1; 9091/1; 9121/1; 9151/1; 9181/1; 9211/1; 9241/1; 9271/1; 9301/1; 9331/1; 9361/1; 9391/1; 9421/1; 9451/1; 9481/1; 9511/1; 9541/1; 9571/1; 9601/1; 9631/1; 9661/1; 9691/1; 9721/1; 9751/1; 9781/1; 9811/1; 9841/1; 9871/1; 9901/1; 9931/1; 9961/1; 9991/1; 1001/1; 1004/1; 1007/1; 1010/1; 1013/1; 1016/1; 1019/1; 1022/1; 1025/1; 1028/1; 1031/1; 1034/1; 1037/1; 1040/1; 1043/1; 1046/1; 1049/1; 1052/1; 1055/1; 1058/1; 1061/1; 1064/1; 1067/1; 1070/1; 1073/1; 1076/1; 1079/1; 1082/1; 1085/1; 1088/1; 1091/1; 1094/1; 1097/1; 1100/1; 1103/1; 1106/1; 1109/1; 1112/1; 1115/1; 1118/1; 1121/1; 1124/1; 1127/1; 1130/1; 1133/1; 1136/1; 1139/1; 1142/1; 1145/1; 1148/1; 1151/1; 1154/1; 1157/1; 1160/1; 1163/1; 1166/1; 1169/1; 1172/1; 1175/1; 1178/1; 1181/1; 1184/1; 1187/1; 1190/1; 1193/1; 1196/1; 1199/1; 1202/1; 1205/1; 1208/1; 1211/1; 1214/1; 1217/1; 1220/1; 1223/1; 1226/1; 1229/1; 1232/1; 1235/1; 1238/1; 1241/1; 1244/1; 1247/1; 1250/1; 1253/1; 1256/1; 1259/1; 1262/1; 1265/1; 1268/1; 1271/1; 1274/1; 1277/1; 1280/1; 1283/1; 1286/1; 1289/1; 1292/1; 1295/1; 1298/1; 1301/1; 1304/1; 1307/1; 1310/1; 1313/1; 1316/1; 1319/1; 1322/1; 1325/1; 1328/1; 1331/1; 1334/1; 1337/1; 1340/1; 1343/1; 1346/1; 1349/1; 1352/1; 1355/1; 1358/1; 1361/1; 1364/1; 1367/1; 1370/1; 1373/1; 1376/1; 1379/1; 1382/1; 1385/1; 1388/1; 1391/1; 1394/1; 1397/1; 1400/1; 1403/1; 1406/1; 1409/1; 1412/1; 1415/1; 1418/1; 1421/1; 1424/1; 1427/1; 1430/1; 1433/1; 1436/1; 1439/1; 1442/1; 1445/1; 1448/1; 1451/1; 1454/1; 1457/1; 1460/1; 1463/1; 1466/1; 1469/1; 1472/1; 1475/1; 1478/1; 1481/1; 1484/1; 1487/1; 1490/1; 1493/1; 1496/1; 1499/1; 1502/1; 1505/1; 1508/1; 1511/1; 1514/1; 1517/1; 1520/1; 1523/1; 1526/1; 1529/1; 1532/1; 1535/1; 1538/1; 1541/1; 1544/1; 1547/1; 1550/1; 1553/1; 1556/1; 1559/1; 1562/1; 1565/1; 1568/1; 1571/1; 1574/1; 1577/1; 1580/1; 1583/1; 1586/1; 1589/1; 1592/1; 1595/1; 1598/1; 1601/1; 1604/1; 1607/1; 1610/1; 1613/1; 1616/1; 1619/1; 1622/1; 1625/1; 1628/1; 1631/1; 1634/1; 1637/1; 1640/1; 1643/1; 1646/1; 1649/1; 1652/1; 1655/1; 1658/1; 1661/1; 1664/1; 1667/1; 1670/1; 1673/1; 1676/1; 1679/1; 1682/1; 1685/1; 1688/1; 1691/1; 1694/1; 1697/1; 1700/1; 1703/1; 1706/1; 1709/1; 1712/1; 1715/1; 1718/1; 1721/1; 1724/1; 1727/1; 1730/1; 1733/1; 1736/1; 1739/1; 1742/1; 1745/1; 1748/1; 1751/1; 1754/1; 1757/1; 1760/1; 1763/1; 1766/1; 1769/1; 1772/1; 1775/1; 1778/1; 1781/1; 1784/1; 1787/1; 1790/1; 1793/1; 1796/1; 1799/1; 1802/1; 1805/1; 1808/1; 1811/1; 1814/1; 1817/1; 1820/1; 1823/1; 1826/1; 1829/1; 1832/1; 1835/1; 1838/1; 1841/1; 1844/1; 1847/1; 1850/1; 1853/1; 1856/1; 1859/1; 1862/1; 1865/1; 1868/1; 1871/1; 1874/1; 1877/1; 1880/1; 1883/1; 1886/1; 1889/1; 1892/1; 1895/1; 1898/1; 1901/1; 1904/1; 1907/1; 1910/1; 1913/1; 1916/1; 1919/1; 1922/1; 1925/1; 1928/1; 1931/1; 1934/1; 1937/1; 1940/1; 1943/1; 1946/1; 1949/1; 1952/1; 1955/1; 1958/1; 1961/1; 1964/1; 1967/1; 1970/1; 1973/1; 1976/1; 1979/1; 1982/1; 1985/1; 1988/1; 1991/1; 1994/1; 1997/1; 2000/1; 2003/1; 2006/1; 2009/1; 2012/1; 2015/1; 2018/1; 2021/1; 2024/1; 2027/1; 2030/1; 2033/1; 2036/1; 2039/1; 2042/1; 2045/1; 2048/1; 2051/1; 2054/1; 2057/1; 2060/1; 2063/1; 2066/1; 2069/1; 2072/1; 2075/1; 2078/1; 2081/1; 2084/1; 2087/1; 2090/1; 2093/1; 2096/1; 2099/1; 2102/1; 2105/1; 2108/1; 2111/1; 2114/1; 2117/1; 2120/1; 2123/1; 2126/1; 2129/1; 2132/1; 2135/1; 2138/1; 2141/1; 2144/1; 2147/1; 2150/1; 2153/1; 2156/1; 2159/1; 2162/1; 2165/1; 2168/1; 2171/1; 2174/1; 2177/1; 2180/1; 2183/1; 2186/1; 2189/1; 2192/1; 2195/1; 2198/1; 2201/1; 2204/1; 2207/1; 2210/1; 2213/1; 2216/1; 2219/1; 2222/1; 2225/1; 2228/1; 2231/1; 2234/1; 2237/1; 2240/1; 2243/1; 2246/1; 2249/1; 2252/1; 2255/1; 2258/1; 2261/1; 2264/1; 2267/1; 2270/1; 2273/1; 2276/1; 2279/1; 2282/1; 2285/1; 2288/1; 2291/1; 2294/1; 2297/1; 2300/1; 2303/1; 2306/1; 2309/1; 2312/1; 2315/1; 2318/1; 2321/1; 2324/1; 2327/1; 2330/1; 2333/1; 2336/1; 2339/1; 2342/1; 2345/1; 2348/1; 2351/1; 2354/1; 2357/1; 2360/1; 2363/1; 2366/1; 2369/1; 2372/1; 2375/1; 2378/1; 2381/1; 2384/1; 2387/1; 2390/1; 2393/1; 2396/1; 2399/1; 2402/1; 2405/1; 2408/1; 2411/1; 2414/1; 2417/1; 2420/1; 2423/1; 2426/1; 2429/1; 2432/1; 2435/1; 2438/1; 2441/1; 2444/1; 2447/1; 2450/1; 2453/1; 2456/1; 2459/1; 2462/1; 2465/1; 2468/1; 2471/1; 2474/1; 2477/1; 2480/1; 2483/1; 2486/1; 2489/1; 2492/1; 2495/1; 2498/1; 2501/1; 2504/1; 2507/1; 2510/1; 2513/1; 2516/1; 2519/1; 2522/1; 2525/1; 2528/1; 2531/1; 2534/1; 2537/1; 2540/1; 2543/1; 2546/1; 2549/1; 2552/1; 2555/1; 2558/1; 2561/1; 2564/1; 2567/1; 2570/1; 2573/1; 2576/1; 2579/1; 2582/1; 2585/1; 2588/1; 2591/1; 2594/1; 2597/1; 2600/1; 2603/1; 2606/1; 2609/1; 2612/1; 2615/1; 2618/1; 2621/1; 2624/1; 2627/1; 2630/1; 2633/1; 2636/1; 2639/1; 2642/1; 2645/1; 2648/1; 2651/1; 2654/1; 2657/1; 2660/1; 2663/1; 2666/1; 2669/1; 2672/1; 2675/1; 2678/1; 2681/1; 2684/1; 2687/1; 2690/1; 2693/1; 2696/1; 2699/1; 2702/1; 2705/1; 2708/1; 2711/1; 2714/1; 2717/1; 2720/1; 2723/1; 2726/1; 2729/1; 2732/1; 2735/1; 2738/1; 2741/1; 2744/1; 2747/1; 2750/1; 2753/1; 2756/1; 2759/1; 2762/1; 2765/1; 2768/1; 2771/1; 2774/1; 2777/1; 2780/1; 2783/1; 2786/1; 2789/1; 2792/1; 2795/1; 2798/1; 2801/1; 2804/1; 2807/1; 2810/1; 2813/1; 2816/1; 2819/1; 2822/1; 2825/1; 2828/1; 2831/1; 2834/1; 2837/1; 2840/1; 2843/1; 2846/1; 2849/1; 2852/1; 2855/1; 2858/1; 2861/1; 2864/1; 2867/1; 2870/1; 2873/1; 2876/1; 2879/1; 2882/1; 2885/1; 2888/1; 2891/1; 2894/1; 2897/1; 2900/1; 2903/1; 2906/1; 2909/1; 2912/1; 2915/1; 2918/1; 2921/1; 2924/1; 2927/1; 2930/1; 2933/1; 2936/1; 2939/1; 2942/1; 2945/1; 2948/1; 2951/1; 2954/1; 2957/1; 2960/1; 2963/1; 2966/1; 2969/1; 2972/1; 2975/1; 2978/1; 2981/1; 2984/1; 2987/1; 2990/1; 2993/1; 2996/1; 2999/1; 3002/1; 3005/1; 3008/1; 3011/1; 3014/1; 3017/1; 3020/1; 3023/1; 3026/1; 3029/1; 3032/1; 3035/1; 3038/1; 3041/1; 3044/1; 3047/1; 3050/1; 3053/1; 3056/1; 3059/1; 3062/1; 3065/1; 3068/1; 3071/1; 3074/1; 3077/1; 3080/1; 3083/1; 3086/1; 3089/1; 3092/1; 3095/1; 3098/1; 3101/1; 3104/1; 3107/1; 3110/1; 3113/1; 3116/1; 3119/1; 3122/1; 3125/1; 3128/1; 3131/1; 3134/1; 3137/1; 3140/1; 3143/1; 3146/1; 3149/1; 3152/1; 3155/1; 3158/1; 3161/1; 3164/1; 3167/1; 3170/1; 3173/1; 3176/1; 3179/1; 3182/1; 3185/1; 3188/1; 3191/1; 3194/1; 3197/1; 3200/1; 3203/1; 3206/1; 3209/1; 3212/1; 3215/1; 3218/1; 3221/1; 3224/1; 3227/1; 3230/1; 3233/1; 3236/1; 3239/1; 3242/1; 3245/1; 3248/1; 3251/1; 3254/1; 3257/1; 3260/1; 3263/1; 3266/1; 3269/1; 3272/1; 3275/1; 3278/1; 3281/1; 3284/1; 3287/1; 3290/1; 3293/1; 3296/1; 3299/1; 3302/1; 3305/1; 3308/1; 3311/1; 3314/1; 3317/1; 3320/1; 3323/1; 3326/1; 3329/1; 3332/1; 3335/1; 3338/1; 3341/1; 3344/1; 3347/1; 3350/1; 3353/1; 3356/1; 3359/1; 3362/1; 3365/1; 3368/1; 3371/1; 3374/1; 3377/1; 3380/1; 3383/1; 3386/1; 3389/1; 3392/1; 3395/1; 3398/1; 3401/1; 3404/1; 3407/1; 3410/1; 3413/1; 3416/1; 3419/1; 3422/1; 3425/1; 3428/1; 3431/1; 3434/1; 3437/1; 3440/1; 3443/1; 3446/1; 3449/1; 3452/1; 3455/1; 3458/1; 3461/1; 3464/1; 3467/1; 3470/1; 3473/1; 3476/1; 3479/1; 3482/1; 3485/1; 3488/1; 3491/1; 3494/1; 3497/1; 3500/1; 3503/1; 3506/1; 3509/1; 3512/1; 3515/1; 3518/1; 3521/1; 3524/1; 3527/1; 3530/1; 3533/1; 3536/1; 3539/1; 3542/1; 3545/1; 3548/1; 3551/1; 3554/1; 3557/1; 3560/1; 3563/1; 3566/1; 3569/1; 3572/1; 3575/1; 3578/1; 3581/1; 3584/1; 3587/1; 3590/1; 3593/1; 3596/1; 3599/1; 3602/1; 3605/1; 3608/1; 3611/1; 3614/1; 3617/1; 3620/1; 3623/1; 3626/1; 3629/1; 3632/1; 3635/1; 3638/1; 3641/1; 3644/1; 3647/1; 3650/1; 3653/1; 3656/1; 3659/1; 3662/1; 3665/1; 3668/1; 3671/1; 3674/1; 3677/1; 3680/1; 3683/1; 3686/1; 3689/1; 3692/1; 3695/1; 3698/1; 3701/1; 3704/1; 3707/1; 3710/1; 3713/1; 3716/1; 3719/1; 3722/1; 3725/1; 3728/1; 3731/1; 3734/1; 3737/1; 3740/1; 3743/1; 3746/1; 3749/1; 3752/1; 3755/1; 3758/1; 3761/1; 3764/1; 3767/1; 3770/1; 3773/1; 3776/1; 3779/1; 3782/1; 3785/1; 3788/1; 3791/1; 3794/1; 3797/1; 3800/1; 3803/1; 3806/1; 3809/1; 3812/1; 3815/1; 3818/1; 3821/1; 3824/1; 3827/1; 3830/1; 3833/1; 3836/1; 3839/1; 3842/1; 3845/1; 3848/1; 3851/1; 3854/1; 3857/1; 3860/1; 3863/1; 3866/1; 3869/1; 3872/1; 3875/1; 3878/1; 3881/1; 3884/1; 3887/1; 3890/1; 3893/1; 3896/1; 3899/1; 3902/1; 3905/1; 3908/1; 3911/1; 3914/1; 3917/1; 3920/1; 3923/1; 3926/1; 3929/1; 3932/1; 3935/1; 3938/1; 3941/1; 3944/1; 3947/1; 3950/1; 3953/1; 3956/1; 3959/1; 3962/1; 3965/1; 3968/1; 3971/1; 3974/1; 3977/1; 3980/1; 3983/1; 3986/1; 3989/1; 3992/1; 3995/1; 3998/1; 4001/1; 4004/1; 4007/1; 4010/1; 4013/1; 4016/1; 4019/1; 4022/1; 4025/1; 4028/1; 4031/1; 4034/1; 4037/1; 4040/1; 4043/1; 4046/1; 4049/1; 4052/1; 4055/1; 4058/1; 4061/1; 4064/1; 4067/1; 4070/1; 4073/1; 4076/1; 4079/1; 4082/1; 4085/1; 4088/1; 4091/1; 4094/1; 4097/1; 4100/1; 4103/1; 4106/1; 4109/1; 4112/1; 4115/1; 4118/1; 4121/1; 4124/1; 4127/1; 4130/1; 4133/1; 4136/1; 4139/1; 4142/1; 4145/1; 4148/1; 4151/1; 4154/1; 4157/1; 4160/1; 4163/1; 4166/1; 4169/1; 4172/1; 4175/1; 4178/1; 4181/1; 4184/1; 4187/1; 4190/1; 4193/1; 4196/1; 4199/1; 4202/1; 4205/1; 4208/1; 4211/1; 4214/1; 4217/1; 4220/1; 4223/1; 4226/1; 4229/1; 4232/1; 4235/1; 4238/1; 4241/1; 4244/1; 4247/1; 4250/1; 4253/1; 4256/1; 4259/1; 4262/1; 4265/1; 4268/1; 4271/1; 4274/1; 4277/1; 4280/1; 4283/1; 4286/1; 4289/1; 4292/1; 4295/1; 4298/1; 4301/1; 4304/1; 4307/1; 4310/1; 4313/1; 4316/1; 4319/1; 4322/1; 4325/1; 4328/1; 4331/1; 4334/1; 4337/1; 4340/1; 4343/1; 4346/1; 4349/1; 4352/1; 4355/1; 4358/1; 4361/1; 4364/1; 4367/1; 4370/1; 4373/1; 4376/1; 4379/1; 4382/1; 4385/1; 4388/1; 4391/1; 4394/1; 4397/1; 4400/1; 4403/1; 4406/1; 4409/1; 4412/1; 4415/1; 4418/1; 4421/1; 4424/1; 4427/1; 4430/1; 4433/1; 4436/1; 4439/1; 4442/1; 4445/1; 4448/1; 4451/1; 4454/1; 4457/1; 4460/1; 4463/1; 4466/1; 4469/1; 4472/1; 4475/1; 4478/1; 4481/1; 4484/1; 4487/1; 4490/1; 4493/1; 4496/1; 4499/1; 4502/1; 4505/1; 4508/1; 4511/1; 4514/1; 4517/1; 4520/1; 4523/1; 4526/1; 4529/1; 4532/1; 4535/1; 4538/1; 4541/1; 4544/1; 4547/1; 4550/1; 4553/1; 4556/1; 4559/1; 4562/1; 4565/1; 4568/1; 4571/

F;1149-1180/Domain: EGF homology <EG17>
F;1187-1218/Domain: EGF homology <EG18>
F;1233-1264/Domain: EGF homology <EG19>
F;1352-1383/Domain: EGF homology <EG19>
F;1391-1425/Domain: EGF homology <EG19>
F;1917-1948/Domain: ankyrin repeat homology <AN1>
F;1949-1981/Domain: ankyrin repeat homology <AN2>
F;1983-2015/Domain: ankyrin repeat homology <AN3>
F;2016-2048/Domain: ankyrin repeat homology <AN4>
F;2049-2081/Domain: ankyrin repeat homology <AN5>

Query Match 12.7%; Score 95.5; DB 2; Length 2531;
Best Local Similarity 26.8%; Pred. No. 0.49;
Matches 34; Conservative 7; Mismatches 35; Indels 51; Gaps 7;

QY 15 ISACAMLLCHGSLQHTFQHHLRPEGTCT-VIAAHR-----CNKN 56
Db 679 IDECAGSPCH-----NGTCDGAGTCRCPEGYHPTCLSEVNECNSN 723
QY 57 -----RIERSQTVKSCLPKGVAGTTRNRPSCVDASIVIGKWCMEPCLEGECKTLP 111
Db 724 PCIHGACRDGLNGYKCDAPG-WSGTNCD-----INNNECESNPVCVNGGTCKDM- 771
QY 112 DNSGWMC 118
Db 772 -TSGYVC 777

RESULT 5
T31070
Notch homolog protein - sea squirt (Halocynthia roretzi)
C;Species: Halocynthia roretzi
C;Date: 02-Sep-2000 #sequence_revision 02-Sep-2000 #text_change 08-Sep-2002
C;Accession: T31070
R;Hori, S.; Saitoh, T.; Matsumoto, M.; Makabe, K.W.; Nishida, H.
Dev. Genes Evol. 207, 371-380, 1997
A;Title: Notch homologue from Halocynthia roretzi is preferentially expressed in the central nervous system
A;Reference number: Z20775
A;Accession: T31070
A;Status: preliminary; translated from GB/EMBL/DBJ
A;Molecule type: mRNA
A;Residues: 1-2352 <HOR>
A;Cross-references: EMBL:AB001327; NID:di204472; PID:di026501; PIDN:BA25571.1
C;Genetics:
A;Gene: Notch
C;Superfamily: notch protein; ankyrin repeat homology; EGF homology

Query Match 12.5%; Score 94.5; DB 2; Length 2352;
Best Local Similarity 26.0%; Pred. No. 0.57;
Matches 27; Conservative 12; Mismatches 34; Indels 31; Gaps 5;

QY 32 QQHHLRPEGG-TCEVIAAH-----RCCKNRIERSQTVKSCLPKQVA 75
Db 236 KRRDLQTEGGFTCNVYGFTRDDCSENIIDCSNVACFHNARCIDQAGTFECLCTPG--- 292
QY 76 GTTRNRPSC-VDASIVIGKWCMEPCLEGECKTLPDNSGWMC 118
Db 293 ---NRILCHLDDA-----CISDPGARGATCDTNPITGHWC 325

RESULT 6
T31070
notch homolog - sea urchin (Lytechinus variegatus)
C;Species: Lytechinus variegatus (variegated urchin)
C;Date: 22-Oct-1999 #sequence_revision 22-Oct-1999 #text_change 31-Jan-2000
C;Accession: T31070
R;Sherwood, D.R.; McClay, D.R.
Development 124, 3363-3374, 1997
A;Title: Identification and localization of a sea urchin Notch homologue: insights into the evolution of the Notch signaling pathway
A;Reference number: Z20966; MUID:97454256; PMID:9310331
A;Accession: T31070
A;Status: preliminary; translated from GB/EMBL/DBJ
A;Molecule type: mRNA

F;1149-1180/Domain: EGF homology <EG17>
F;1187-1218/Domain: EGF homology <EG18>
F;1233-1264/Domain: EGF homology <EG19>
F;1352-1383/Domain: EGF homology <EG19>
F;1391-1425/Domain: EGF homology <EG19>
F;1917-1948/Domain: ankyrin repeat homology <AN1>
F;1949-1981/Domain: ankyrin repeat homology <AN2>
F;1983-2015/Domain: ankyrin repeat homology <AN3>
F;2016-2048/Domain: ankyrin repeat homology <AN4>
F;2049-2081/Domain: ankyrin repeat homology <AN5>

Query Match 12.7%; Score 95.5; DB 2; Length 2531;
Best Local Similarity 26.8%; Pred. No. 0.49;
Matches 34; Conservative 7; Mismatches 35; Indels 51; Gaps 7;

QY 15 ISACAMLLCHGSLQHTFQHHLRPEGTCT-VIAAHR-----CNKN 56
Db 679 IDECAGSPCH-----NGTCDGAGTCRCPEGYHPTCLSEVNECNSN 723
QY 57 -----RIERSQTVKSCLPKGVAGTTRNRPSCVDASIVIGKWCMEPCLEGECKTLP 111
Db 724 PCIHGACRDGLNGYKCDAPG-WSGTNCD-----INNNECESNPVCVNGGTCKDM- 771
QY 112 DNSGWMC 118
Db 772 -TSGYVC 777

RESULT 5
T31070
Notch homolog protein - sea squirt (Halocynthia roretzi)
C;Species: Halocynthia roretzi
C;Date: 02-Sep-2000 #sequence_revision 02-Sep-2000 #text_change 08-Sep-2002
C;Accession: T31070
R;Hori, S.; Saitoh, T.; Matsumoto, M.; Makabe, K.W.; Nishida, H.
Dev. Genes Evol. 207, 371-380, 1997
A;Title: Notch homologue from Halocynthia roretzi is preferentially expressed in the central nervous system
A;Reference number: Z20775
A;Accession: T31070
A;Status: preliminary; translated from GB/EMBL/DBJ
A;Molecule type: mRNA
A;Residues: 1-2352 <HOR>
A;Cross-references: EMBL:AB001327; NID:di204472; PID:di026501; PIDN:BA25571.1
C;Genetics:
A;Gene: Notch
C;Superfamily: notch protein; ankyrin repeat homology; EGF homology

Query Match 12.5%; Score 94.5; DB 2; Length 2352;
Best Local Similarity 26.0%; Pred. No. 0.57;
Matches 27; Conservative 12; Mismatches 34; Indels 31; Gaps 5;

QY 32 QQHHLRPEGG-TCEVIAAH-----RCCKNRIERSQTVKSCLPKQVA 75
Db 236 KRRDLQTEGGFTCNVYGFTRDDCSENIIDCSNVACFHNARCIDQAGTFECLCTPG--- 292
QY 76 GTTRNRPSC-VDASIVIGKWCMEPCLEGECKTLPDNSGWMC 118
Db 293 ---NRILCHLDDA-----CISDPGARGATCDTNPITGHWC 325

RESULT 6
T31070
notch homolog - sea urchin (Lytechinus variegatus)
C;Species: Lytechinus variegatus (variegated urchin)
C;Date: 22-Oct-1999 #sequence_revision 22-Oct-1999 #text_change 31-Jan-2000
C;Accession: T31070
R;Sherwood, D.R.; McClay, D.R.
Development 124, 3363-3374, 1997
A;Title: Identification and localization of a sea urchin Notch homologue: insights into the evolution of the Notch signaling pathway
A;Reference number: Z20966; MUID:97454256; PMID:9310331
A;Accession: T31070
A;Status: preliminary; translated from GB/EMBL/DBJ
A;Molecule type: mRNA

A;Residues: 1-2531 <SHE>
A;Cross-references: EMBL:AF000634; NID:92570350; PID:92570351; PIDN:AA82088.1
C;Superfamily: notch protein; ankyrin repeat homology; EGF homology

Query Match 12.3%; Score 93; DB 2; Length 2531;
Best Local Similarity 22.4%; Pred. No. 0.85;
Matches 28; Conservative 18; Mismatches 41; Indels 38; Gaps 5;

QY 40 EGGTCEV-----IAAHRCCNKNRIERSQTVKSCLPKGVAGTTRNRPSCV----- 85
Db 350 EGGTCEINKDDCTPNPCQFEGCEDRVASPKCTCPGRTGLLCHLEDACMSNCPCHTAQC 409
QY 86 DASIVIGKWCMEPCLEGECKTLPD-----NSGWMCA-----TG 121
Db 410 STSVWDGSGFTCDATGYQGFNCSEDIDECSSLMSDICSQSGTCTCNFGDGSCLCSGFTG 469
QY 122 NKIKT 126
Db 470 SRET 474

RESULT 7
A35672
crumbs protein - fruit fly (Drosophila melanogaster)
C;Species: Drosophila melanogaster
C;Date: 21-Sep-1990 #sequence_revision 18-Nov-1992 #text_change 21-Jul-2003
C;Accession: A35672
R;Tepass, U.; Theres, C.; Knust, E.
Cell 61, 787-799, 1990
A;Title: crumbs encodes an EGF-like protein expressed on apical membranes of Drosophila epithelial cells
A;Reference number: A35672; MUID:90263104; PMID:2344615
A;Accession: A35672
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-2139 <TEP>
A;Cross-references: GB:M33753
A;Note: the authors translated the codon GGC for residue 1928 as Cys, and TAT for residue 1929 as Phe
C;Genetics:
A;Gene: FlyBase:crb
A;Cross-references: FlyBase:FBgn0000368
C;Keywords: transmembrane protein
F;352-385/Domain: EGF homology <EGX1>
F;392-424/Domain: EGF homology <EGF1>
F;691-722/Domain: EGF homology <EGF>
F;767-799/Domain: EGF homology <EGF3>
F;1878-1914/Domain: EGF homology <EGX2>

Query Match 12.3%; Score 92.5; DB 2; Length 2139;
Best Local Similarity 25.0%; Pred. No. 0.82;
Matches 27; Conservative 14; Mismatches 46; Indels 17; Gaps 5;

QY 41 GGTCEVIAAHRCC-----NKNRIERSQTVKSCLPKGVAGTTRNRPSCVDASIVIGKWC 96
Db 559 GGTCSMNGTHCYCAVGYSGDRCEKAENCSPLNCPQEPWVC--VQNCCLCPENKVCN---QC 613
QY 97 EMEPCLEGECKTLPDNS-----GWCAT-GNKIKTTIHRP 132
Db 614 ATPCQNGGECVDLPNGDYECCKTRGTGTCGNDVDDECTLHPK 657

RESULT 8
A24420
notch protein - fruit fly (Drosophila melanogaster)
N;Alternate names: neurogenic repetitive locus protein
C;Species: Drosophila melanogaster
C;Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 10-Sep-1999
C;Accession: A24420; A24768; S09358; A05267
R;Kidd, S.; Kelley, M.R.; Young, M.W.
Mol. Cell. Biol. 6, 3094-3108, 1986
A;Reference number: A24420; MUID:87064624; PMID:3097517
A;Accession: A24420
A;Molecule type: DNA
A;Residues: 1-2703 <KID>

A;Cross-references: GB:K03508; NID:G157991; PIDN:AAA28725.1; PID:G157993
 R;Wharton, K.A.; Johansen, K.M.; Xu, T.; Artavanis-Tsakonas, S.
 Cell 43, 567-581, 1985
 A;Reference number: A24768; MUID:86079539; PMID:3935325
 A;Accession: A24768
 A;Molecule type: mRNA
 A;Residues: 1-48, 'I', 50-118, 'R', 120-230, 'I', 232-256, 'N', 258-266, 'A', 268-872, 'R', 874-958,
 A;Note: the authors translated the codon ATC for residue 49 as Thr, ATT for residue 2044
 R;Tautz, D.
 Nucleic Acids Res. 17, 6463-6471, 1989
 A;Title: Hypervariability of simple sequences as a general source for polymorphic DNA ma
 A;Reference number: S09358; MUID:89385974; PMID:2780284
 A;Accession: S09358
 A;Molecule type: DNA
 A;Residues: 2505-2551, 'QQQ', 2552-2576, 'E', 2578-2604 'TAU'
 R;Wharton, K.A.; Yedvobnick, B.; Finnerty, V.G.; Artavanis-Tsakonas, S.
 Cell 40, 55-62, 1985
 A;Title: opa: a novel family of transcribed repeats shared by the Notch locus and other
 A;Reference number: A05267; MUID:85099329; PMID:2981631
 A;Accession: A05267
 A;Molecule type: DNA
 A;Residues: 2504-2576, 'E', 2578-2611 'WHA2'
 C;Genetics:
 A;Gene: notch; opa
 A;Cross-references: FlyBase:FBgn0004647
 A;Map position: 8.96-9.36
 A;Intons: 53/3; 84/3; 171/3; 240/3; 283/3; 2333/3; 2436/3; 2588/3
 C;Superfamily: notch protein; ankyrin repeat homology; EGF homology
 C;Keywords: differentiation; tandem repeat; transmembrane protein
 F;27-43/Domain: transmembrane #status predicted <TMW1>
 F;297-328/Domain: EGF homology <EGX1>
 F;530-561/Domain: EGF homology <EGF1>
 F;568-599/Domain: EGF homology <EGF>
 F;988-1019/Domain: EGF homology <EGX2>
 F;1064-1095/Domain: EGF homology <EGF3>
 F;1187-1218/Domain: EGF homology <EGX3>
 F;1746-1762/Domain: transmembrane #status predicted <TMW2>
 F;1950-1982/Domain: ankyrin repeat homology <AN1>
 F;1983-2015/Domain: ankyrin repeat homology <AN2>
 F;1988-2004/Domain: transmembrane #status predicted <TMW3>
 F;2017-2049/Domain: ankyrin repeat homology <AN3>
 F;2050-2082/Domain: ankyrin repeat homology <AN4>
 F;2083-2115/Domain: ankyrin repeat homology <AN5>
 F;2538-2568/Region: glutamine-rich
 F;2538-2568/Domain: neurogenic repetitive element #status predicted <OPA>
 Query Match 11.5%; Score 86.5; DB 1; Length 2703;
 Best Local Similarity 25.9%; Pred. No. 3.9;
 Matches 22; Conservative 10; Mismatches 34; Indels 19; Gaps 3;
 QY 40 EGGTCEV-----IAHRCCKNRIEERSQTVKCSCLPGKVA-----GTTNRPS 83
 Db 1214 QGONCELINDCAPNQCNGTGRVNFSCPPGTGTGIICEINKDKCKPGACHNGS 1273
 QY 84 CVDASIVIGKWCMEPCLEGECK 108
 Db 1274 CIDR---VGGFECVQPGFVGARCE 1295
 RESULT 9
 JC7570
 Delta-4 protein - human
 C;Species: Homo sapiens (man)
 C;Date: 30-Jun-2001 #sequence_revision 30-Jun-2001 #text_change 16-Aug-2002
 R;Yoneya, T.; Tahara, T.; Nagao, K.; Yamada, Y.; Yamamoto, T.; Osawa, M.; Miyatani, S.;
 J. Biochem. 129, 27-34, 2001
 A;Title: Molecular cloning of Delta-4, a new mouse and human notch ligand.
 A;Reference number: JC7569; MUID: 21064937; PMID:11134954
 A;Accession: JC7570
 A;Molecule type: mRNA
 A;Residues: 1-685 <YON>
 A;Cross-references: DDBJ:AB043894

C;Comment: This protein, a member of the Notch family of proteins, is a transmembrane re
 ates the Notch signaling, the growth or differentiation of vascular endothelial cells.
 C;Genetics:
 A;Gene: delta-4
 C;Superfamily: delta-4 protein; EGF homology
 C;Keywords: transmembrane protein
 Query Match 11.4%; Score 86; DB 2; Length 685;
 Best Local Similarity 25.0%; Pred. No. 1.2;
 Matches 28; Conservative 9; Mismatches 45; Indels 30; Gaps 6;
 QY 14 WISAC-----AMLLCHSLOHTFOOHLHREPGTCEVIAHRCCKNRIEERSQTVKCS 69
 Db 268 WQCTDEGWLFCDDQDLYNC-----THSPCKNGATC-----SNSGQRSYT--CTC 312
 QY 70 LPQKVAAGTTNRPSCVDAISIVIGKWCMEPCLEGECKTLPDNSGMCATG 121
 Db 313 RPYVTG-----VDCELEISE--CDNPNCRNGSGCKDQEDGYHCLPPG 353
 RESULT 10
 A56136
 Jagged protein precursor - rat
 C;Species: Rattus norvegicus (Norway rat)
 C;Date: 28-Apr-1995 #sequence_revision 28-Apr-1995 #text_change 21-Jul-2003
 C;Accession: A56136
 R;Kindell, C.E.; Shawber, C.J.; Boulter, J.; Weinmaster, G.
 Cell 80, 909-917, 1995
 A;Title: Jagged: a mammalian ligand that activates Notchl.
 A;Reference number: A56136; MUID:95211842; PMID:7697721
 A;Accession: A56136
 A;Status: preliminary
 A;Molecule type: mRNA
 A;Residues: 1-1220 <LIN>
 A;Cross-references: GB:I38483
 F;379-410/Domain: EGF homology <EGF1>
 F;492-523/Domain: EGF homology <EGF>
 F;634-665/Domain: EGF homology <EGF2>
 Query Match 11.4%; Score 86; DB 2; Length 1220;
 Best Local Similarity 27.0%; Pred. No. 2.1;
 Matches 33; Conservative 9; Mismatches 34; Indels 46; Gaps 9;
 QY 39 PE---GCTCEVIAHRC-----C-NKNRIEERSQTVKCSCLPGKVAAGTTNR 81
 Db 326 PEGYSGPNCE-IAEHACLSDPCHNRGSKETSGFECSPGWTGPTCTNIDDCSPNNC 384
 QY 82 ---PSCVDASIVIG-----KWW-----CEMEPCLEGECKTLPD-----SG 115
 Db 385 SHGGTQD--LVNGFKVCVPPTGKTCTCOLDANECEAKFCVNARSKNLIASYCDCLPG 442
 QY 116 WM 117
 Db 443 WM 444
 RESULT 11
 A40136
 fibropellin Ia - sea urchin (Strongylocentrotus purpuratus)
 N;Alternate names: epidermal growth factor homolog precursor
 N;Contains: alternatively spliced fibropellin Ib (EGFI)
 C;Species: Strongylocentrotus purpuratus (purple urchin)
 C;Date: 13-May-1992 #sequence_revision 17-Sep-1997 #text_change 21-Jul-2000
 C;Accession: A40136; B40136; A40136; A29316; A43131
 R;Delgadillo-Reynoso, M.G.; Rollo, D.R.; Hursh, D.A.; Raff, R.A.
 J. Mol. Evol. 29, 314-327, 1989
 A;Title: Structural analysis of the uEGF gene in the sea urchin Strongylocentrotus purpu
 A;Reference number: A40136; MUID:90112459; PMID:2514273
 A;Accession: A40136
 A;Status: preliminary
 A;Molecule type: mRNA
 A;Residues: 1-114
 A;Cross-references: GB:X17530; NID:g10225; PID:g667061

A:Accession: B40136
A:Status: preliminary; not compared with conceptual translation
A:Molecule type: DNA
A:Residues: 181-251,329-370,'R',372-408,'RA',411-441 <DE2>
A:Accession: C40136
A:Status: preliminary; not compared with conceptual translation
A:Molecule type: DNA
A:Residues: 'K',747-821,898-978 <DE3>
R:Hursh, D.A.; Andrews, M.E.; Raff, R.A.
Science 237, 1487-1490, 1987
A:Title: A sea urchin gene encodes a polypeptide homologous to epidermal growth factor.
A:Reference number: A29316; MUID:87319677; PMID:3498216
A:Accession: A29316
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 'S',280-481,786-1064 <HUR>
A:Cross-references: GB:M17421; NID:G161474; PID:G552260
R:Hunt, L.T.; Barker, W.C.
FASEB J. 3, 1760-1764, 1989
A:Title: Avidin-like domain in an epidermal growth factor homolog from a sea urchin.
A:Reference number: A43131; MUID:89196806; PMID:2784773
A:Contents: annotation
C:Comment: EGF homology repeats 10-17 are spliced out in the short form (fibropellin Ib)
C:Superfamily: C1r/C1s repeat homology; EGF homology
F:1-19/Domain: signal sequence #status predicted <SIG>
F:20-1064/Product: fibropellin I #status predicted <FIB>
F:23-54/Domain: EGF homology <EG01>
F:57-175/Domain: C1r/C1s repeat homology <C1r>
F:180-211/Domain: EGF homology <EG02>
F:218-249/Domain: EGF homology <EG03>
F:256-287/Domain: EGF homology <EG04>
F:294-325/Domain: EGF homology <EG05>
F:332-363/Domain: EGF homology <EG06>
F:370-401/Domain: EGF homology <EG07>
F:408-439/Domain: EGF homology <EG08>
F:446-477/Domain: EGF homology <EG09>
F:484-515/Domain: EGF homology <EG10>
F:522-553/Domain: EGF homology <EG11>
F:560-591/Domain: EGF homology <EG12>
F:598-629/Domain: EGF homology <EG13>
F:636-667/Domain: EGF homology <EG14>
F:674-705/Domain: EGF homology <EG15>
F:712-743/Domain: EGF homology <EG16>
F:750-781/Domain: EGF homology <EG17>
F:788-819/Domain: EGF homology <EG18>
F:826-857/Domain: EGF homology <EG19>
F:864-895/Domain: EGF homology <EG20>
F:902-933/Domain: EGF homology <EG21>
F:936-1064/Region: avidin-like
F:23-34,28-43,45-54,62-88,180-191,185-200,202-211,218-229,223-238,240-249,256-267,261-275,451-466,468-477,484-495/Disulfide bonds: #status predicted
F:489-504,506-515,522-533,527-542,544-553,560-571,565-580,582-591,598-609,603-618,620-624,810-819,826-837,831-846,848-857,864-875,869-884,886-895,902-913,907-922,924-933/Disul

Query Match 11.3%; Score 85.5; DB 2; Length 1064;
Best Local Similarity 28.1%; Pred. No. 2;
Matches 25; Conservative 9; Mismatches 34; Indels 21; Gaps 4;
QY 40 EGGTCEV-----IAAHRCCNKNRIERSQTVKSCLPKGV---AGTTNRNPSQVDASIVIG 92
DB 625 EGTNCEINTDECASSPQWGLCVQVNSVCFCLPGFTGIHCGTE-----IDE----- 673
QY 93 KWCMEPCLEGECKTLDPNCGWMCATG 121
DB 674 ---CASSPLNGGQCITDRVDSVEVCVCAAG 699
RESULT 12
A35844
Xotch protein - African clawed frog
C:Species: Xenopus laevis (African clawed frog)
C:Date: 12-Oct-1990 #sequence_revision 12-Oct-1990 #text_change 02-Aug-2002
C:Accession: A35844

R:Coffman, C.; Harris, W.; Kintner, C.
Science 249, 1438-1441, 1990
A:Title: Xotch, the Xenopus homolog of Drosophila notch.
A:Reference number: A35844; MUID:90385285; PMID:2402639
A:Accession: A35844
A:Status: preliminary; nucleic acid sequence not shown; not compared with conceptual translation
A:Molecule type: mRNA
A:Residues: 1-2524 <COF>
C:Superfamily: notch protein; ankyrin repeat homology; EGF homology
C:Keywords: transmembrane protein
F:146-177/Domain: EGF homology <EGX1>
F:184-215/Domain: EGF homology <EGF1>
F:222-254/Domain: EGF homology <EGF>
F:456-487/Domain: EGF homology <EGX2>
F:757-788/Domain: EGF homology <EGF3>
F:1025-1056/Domain: EGF homology <EGX3>
F:1924-1956/Domain: ankyrin repeat homology <AN1>
F:1957-1989/Domain: ankyrin repeat homology <AN2>
F:1931-2023/Domain: ankyrin repeat homology <AN3>
F:2024-2056/Domain: ankyrin repeat homology <AN4>
F:2057-2089/Domain: ankyrin repeat homology <AN5>
Query Match 11.1%; Score 84; DB 2; Length 2524;
Best Local Similarity 23.4%; Pred. No. 6.4;
Matches 29; Conservative 16; Mismatches 41; Indels 38; Gaps 6;
QY 23 CHGSLQHTFOQHILRPFGTCEVI-----AAHRCNKNRIERS--QTVKSCLPKGVAG 76
DB 815 CNCMLPVT-----GAICEAVLAPCAGSPCKNGRCCKEDEFETSCPCPGWQGG 864
QY 77 T-----TNRNPSQVDASIVIGKWCMEPCLEGE-----EECKTLDPNCGWM 117
DB 865 TCEIDMNECVNRPCRNGATCONTN---GSYKCNCKPGYTGRTNCEMDIIDCCPNPCHNGS 921
QY 118 CATG 121
DB 922 CSDG 925
RESULT 13
A49175
Notch B protein - mouse (fragment)
N:Alternate names: Notch homolog
C:Species: Mus musculus (house mouse)
C:Date: 21-Jan-1994 #sequence_revision 05-Jan-1996 #text_change 08-Sep-2002
C:Accession: A49175; PH1570; S32113
R:Lardelli, M.; Lendahl, U.
Exp. Cell Res. 204, 364-372, 1993
A:Title: Notch A and Notch B--two mouse Notch homologues coexpressed in a wide variety of tissues.
A:Reference number: A49175; MUID:93178563; PMID:8440332
A:Accession: A49175
A:Status: preliminary; nucleic acid sequence not shown
A:Molecule type: mRNA
A:Residues: 1-1203 <LAR>
A:Cross-references: EMBL:X68279; NID:G287989; PIDN:CAA48340.1; PID:G287990
A:Experimental source: embryo
A:Note: sequence extracted from NCBI backbone (NCBIP:126158)
C:Comment: This protein has many EGF repeats and lin-12/Notch repeats.
C:Superfamily: notch protein; ankyrin repeat homology; EGF homology
F:143-174/Domain: EGF homology <EGX1>
F:482-513/Domain: EGF homology <EGF1>
F:560-591/Domain: EGF homology <EGF>
F:674-705/Domain: EGF homology <EGX2>
F:712-743/Domain: EGF homology <EGF3>
F:836-867/Domain: EGF homology <EGX3>
Query Match 10.9%; Score 82.5; DB 2; Length 1203;
Best Local Similarity 26.4%; Pred. No. 4.5;
Matches 23; Conservative 11; Mismatches 36; Indels 17; Gaps 4;
QY 40 EGGTCEV---IAAHRCCNKNRIERSQTVKSCLPKGVAGTTRNPSQ-VDSIVIGKW 94

Db 170 KGVHCELVNECQSNPCVNNNGQVCKVNRFCQLCPGFTG-----PVCQIDIDD----- 218

QY 95 WCEMEPCLEGECKTLDPDNGWMCATG 121

Db 219 -CSSTPCLINGAKCIDHPNGYECQCQATG 244

RESULT 14

A49128

cell-fate determining gene Notch2 protein - rat

C;Species: Rattus norvegicus (Norway rat)

C;Date: 21-Jan-1994 #sequence_revision 18-Nov-1994 #text_change 02-Aug-2002

C;Accession: A49128

R;Weinmaster, G.; Roberts, V.J.; Lemke, G.

Development 116, 931-941, 1992

A;Title: Notch2: a second mammalian Notch gene.

A;Reference number: A49128; MUID:93202015; PMID:1295745

A;Accession: A49128

A;Status: preliminary; not compared with conceptual translation

A;Molecule type: mRNA

A;Residues: 1-2471 <WEI>

A;Experimental source: Schwann cell

A;Note: sequence extracted from NCBI backbone (NCBIP:127811)

C;Superfamily: notch protein; ankyrin repeat homology; EGF homology

F;264-295/Domain: EGF homology <EGX1>

F;799-830/Domain: EGF homology <EGF1>

F;877-908/Domain: EGF homology <EGX2>

F;1029-1060/Domain: EGF homology <EGF>

F;1067-1098/Domain: EGF homology <EGX3>

F;1153-1184/Domain: EGF homology <EGF3>

F;1191-1222/Domain: EGF homology <EGX4>

F;1876-1908/Domain: ankyrin repeat homology <AN1>

F;1909-1941/Domain: ankyrin repeat homology <AN2>

F;1943-1975/Domain: ankyrin repeat homology <AN3>

F;1976-2008/Domain: ankyrin repeat homology <AN4>

F;2009-2041/Domain: ankyrin repeat homology <AN5>

Query Match

Best Local Similarity 10.9%; Score 82.5; DB 2; Length 2471;

Matches 23; Conservative 11; Mismatches 36; Indels 17; Gaps 4;

QY 40 EGGTCEV-----IAAHRCCNKNRIBERSQTVKCSCLPGKAVGTTNRPS-CVDASIVIGKW 94

Db 487 KGVHCELVNECQSNPCVNNNGQVCKVNRFCQLCPGFTG-----PVCQIDIDD----- 535

QY 95 WCEMEPCLEGECKTLDPDNGWMCATG 121

Db 536 -CSSTPCLINGAKCIDHPNGYECQCQATG 561

RESULT 15

S45306

notch 3 protein - mouse

C;Species: Mus musculus (house mouse)

C;Date: 20-Feb-1995 #sequence_revision 20-Feb-1995 #text_change 02-Aug-2002

C;Accession: S45306

R;Lardelli, M.; Dahlstrand, J.; Lendahl, U.

Mech. Dev. 46, 123-136, 1994

A;Title: The novel Notch homologue mouse Notch 3 lacks specific epidermal growth factor-

A;Reference number: S45306; MUID:95001556; PMID:7918097

A;Accession: S45306

A;Status: preliminary

A;Molecule type: mRNA

A;Residues: 1-2318 <LAR>

A;Cross-references: EMBL:X74760; NID:g483580; PIDN:CAA52776.1; PTD:g483581

C;Superfamily: notch protein; ankyrin repeat homology; EGF homology

F;163-195/Domain: EGF homology <EGF1>

F;474-505/Domain: EGF homology <EGF>

F;854-885/Domain: EGF homology <EGF2>

F;1839-1871/Domain: ankyrin repeat homology <AN1>

F;1872-1904/Domain: ankyrin repeat homology <AN2>

F;1906-1938/Domain: ankyrin repeat homology <AN3>

F;1939-1971/Domain: ankyrin repeat homology <AN4>

F;1972-2004/Domain: ankyrin repeat homology <AN5>

Query Match 10.9%; Score 82; DB 2; Length 2318;

Best Local Similarity 22.8%; Pred. No. 9.2;

Matches 39; Conservative 16; Mismatches 42; Indels 74; Gaps 10;

QY 17 ACAM-----LLCHGSLQHTFQQHHLHRP-----EGTCE-----V 46

Db 340 ACPMGKTGLLCH--LDDACVSNPCHEDAICDTNPFVSGRAICTCPGFTGACDQDVDECS 397

QY 47 IAAHRCNKNRIBERSQTVKCSCLPGK-----VAGTTRNRPSCVDASIVIGKW 94

Db 398 IGANPCHEHLGRCVNTGGSFLCCQCGRGYTGPRCETDVNECLSGPCRQATCLDR---IGQF 454

QY 95 -----WCME-----PCLEGECK-----TLPDN-SGMMC 118

Db 455 TCICMAGFTGTYCEVDIDECQSSPCVNGGVCKDRVNGFSTCFSGFSGSMC 505

Search completed: April 8, 2004, 13:18:05

Job time : 21 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: April 8, 2004, 13:17:35 ; Search time 40 Seconds
(without alignments)
874.269 Million cell updates/sec

Title: US-09-763-335-2

Perfect score: 754

Sequence: 1 MAMVSAMSVMVLYLWISACAM.....SGWMCATGNKIKTRIHPRT 133

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1073127 seqs, 262937947 residues

Total number of hits satisfying chosen parameters: 1073127

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications AA:*

- 1: /cgn2_6/ptodata/1/pubpa/US07_PUBCOMB.pep.*
- 2: /cgn2_6/ptodata/1/pubpa/PCT_NEW_PUB.pep.*
- 3: /cgn2_6/ptodata/1/pubpa/US06_NEW_PUB.pep.*
- 4: /cgn2_6/ptodata/1/pubpa/US06_PUBCOMB.pep.*
- 5: /cgn2_6/ptodata/1/pubpa/US07_NEW_PUB.pep.*
- 6: /cgn2_6/ptodata/1/pubpa/PCTUS_PUBCOMB.pep.*
- 7: /cgn2_6/ptodata/1/pubpa/US08_NEW_PUB.pep.*
- 8: /cgn2_6/ptodata/1/pubpa/US08_PUBCOMB.pep.*
- 9: /cgn2_6/ptodata/1/pubpa/US09A_PUBCOMB.pep.*
- 10: /cgn2_6/ptodata/1/pubpa/US09B_PUBCOMB.pep.*
- 11: /cgn2_6/ptodata/1/pubpa/US09C_PUBCOMB.pep.*
- 12: /cgn2_6/ptodata/1/pubpa/US09D_PUBCOMB.pep.*
- 13: /cgn2_6/ptodata/1/pubpa/US10A_PUBCOMB.pep.*
- 14: /cgn2_6/ptodata/1/pubpa/US10B_PUBCOMB.pep.*
- 15: /cgn2_6/ptodata/1/pubpa/US10C_PUBCOMB.pep.*
- 16: /cgn2_6/ptodata/1/pubpa/US10D_PUBCOMB.pep.*
- 17: /cgn2_6/ptodata/1/pubpa/US10E_PUBCOMB.pep.*
- 18: /cgn2_6/ptodata/1/pubpa/US10F_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	106	14.1	3871	15	US-10-369-493-6538
2	106	14.1	3871	15	US-10-369-493-6538
3	95.5	12.7	2531	12	US-10-072-012-470
4	95.5	12.7	2531	12	US-10-072-012-471
5	95.5	12.7	2531	15	US-10-190-115-29
6	95.5	12.7	2531	15	US-10-369-072-29
7	91.5	12.1	233	9	US-09-216-393-110
8	91.5	12.1	233	14	US-10-321-856-110
9	88.5	11.7	2447	15	US-10-190-115-28
10	88.5	11.7	2447	15	US-10-369-072-28
11	87	11.5	188	9	US-09-764-853-773
12	87	11.5	188	9	US-09-764-898-262
13	87	11.5	188	10	US-09-764-881-109
14	87	11.5	188	14	US-10-073-865-88
15	87	11.5	188	15	US-10-242-747-109

16	86.5	11.5	1139	12	US-10-419-026-4	Sequence 4, Appli
17	86.5	11.5	2556	12	US-10-072-012-134	Sequence 134, App
18	86	11.4	174	9	US-09-764-898-189	Sequence 189, App
19	86	11.4	500	14	US-10-241-476-2	Sequence 2, Appli
20	86	11.4	528	14	US-10-225-630-4	Sequence 4, Appli
21	86	11.4	528	14	US-10-241-476-3	Sequence 3, Appli
22	86	11.4	559	14	US-10-147-493-88	Sequence 88, Appl
23	86	11.4	585	12	US-10-145-127-88	Sequence 88, Appl
24	86	11.4	685	12	US-10-160-503-88	Sequence 88, Appl
25	86	11.4	685	12	US-10-143-118-88	Sequence 88, Appl
26	86	11.4	685	12	US-10-144-993-88	Sequence 88, Appl
27	86	11.4	685	12	US-10-158-787-88	Sequence 88, Appl
28	86	11.4	685	12	US-10-081-056-276	Sequence 276, App
29	86	11.4	685	12	US-10-219-535-214	Sequence 214, App
30	86	11.4	685	12	US-10-232-230-214	Sequence 214, App
31	86	11.4	685	12	US-10-140-024-88	Sequence 88, Appl
32	86	11.4	685	14	US-10-028-072-88	Sequence 88, Appl
33	86	11.4	685	14	US-10-121-049-88	Sequence 88, Appl
34	86	11.4	685	14	US-10-123-904-88	Sequence 88, Appl
35	86	11.4	685	14	US-10-140-470-88	Sequence 88, Appl
36	86	11.4	685	14	US-10-175-746-88	Sequence 88, Appl
37	86	11.4	685	14	US-10-176-918-88	Sequence 88, Appl
38	86	11.4	685	14	US-10-176-921-88	Sequence 88, Appl
39	86	11.4	685	14	US-10-227-884-214	Sequence 214, App
40	86	11.4	685	14	US-10-137-865-88	Sequence 88, Appl
41	86	11.4	685	14	US-10-140-474-88	Sequence 88, Appl
42	86	11.4	685	14	US-10-142-431-88	Sequence 88, Appl
43	86	11.4	685	14	US-10-143-114-88	Sequence 88, Appl
44	86	11.4	685	14	US-10-230-163-214	Sequence 214, App
45	86	11.4	685	14	US-10-140-002-88	Sequence 88, Appl

ALIGNMENTS

RESULT 1

US-10-369-493-6538
; Sequence 6538, Application US/10369493
; Publication No. US20030233675A1
; GENERAL INFORMATION:
; APPLICANT: Cao, Yongwei
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Goldman, Barry S.
; APPLICANT: Chen, Xianfeng
; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
; FILE REFERENCE: 38-10(52052)B
; CURRENT APPLICATION NUMBER: US/10/369,493
; CURRENT FILING DATE: 2003-02-28
; PRIOR APPLICATION NUMBER: US 60/360,039
; PRIOR FILING DATE: 2002-02-21
; NUMBER OF SEQ ID NOS: 47374
; SEQ ID NO 6538
; LENGTH: 3871
; TYPE: PRT
; ORGANISM: Caenorhabditis elegans
US-10-369-493-6538

Query Match 14.1%; Score 106; DB 15; Length 3871;
Best Local Similarity 31.2%; Pred. No. 0.15;
Matches 34; Conservative 13; Mismatches 40; Indels 22; Gaps 6;
QY 24 HGSQHTFQOHUHLRPEGGTCEVIAAHRCKNKNRIERSQ-----TVKCSLPGKVAQT 77
Db 148 HGSLS-----CQYHMSAC-SKTFELCGPHGCHIESIVDPTGSSSDTTTTCICIDWGFVSS 202
QY 78 TRNRSPCVDAISIVIGKWCMECLGECKTLPDN---SGWMCATGNK 123
Db 203 DKNNTCTVDVN-----ECESNFCHPGVDCINLPFGSVCSG--CPKGYK 243

RESULT 2

```

; PRIOR FILING DATE: 2001-01-31
; PRIOR APPLICATION NUMBER: 60/265,412
; PRIOR FILING DATE: 2001-01-31
; PRIOR APPLICATION NUMBER: 60/265,395
; PRIOR FILING DATE: 2001-01-31
; PRIOR APPLICATION NUMBER: 60/266,406
; PRIOR FILING DATE: 2001-02-02
; PRIOR APPLICATION NUMBER: 60/266,767
; PRIOR FILING DATE: 2001-02-05
; PRIOR APPLICATION NUMBER: 60/267,057
; PRIOR FILING DATE: 2001-02-07
; PRIOR APPLICATION NUMBER: 60/266,975
; PRIOR FILING DATE: 2001-02-07
; PRIOR APPLICATION NUMBER: 60/267,459
; PRIOR FILING DATE: 2001-02-08
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 1391
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 470
; LENGTH: 2531
; TYPE: PRT
; ORGANISM: Rattus norvegicus
;
US-10-072-012-470

Query Match          12.7%; Score 95.5; DB 12; Length 2531;
Best Local Similarity 26.8%; Pred. No. 1.2;
Matches 34; Conservative 7; Mismatches 35; Indels 51; Gaps 7;

QY      15 ISACAMLLCHGSLQHTFQOHLHLPGEGTCE-VIAAHC-----CNKN 56
Db      679 IDECAAGSPCH-----NGGTCEDGAGTCTCPGEGYHDPCLSEVNECNSN 723

QY      57 -----RIERSQTVKSGCLPGKVAGTTRNRPSCVDASIVIGKWWCEMEPCLEGECEKTLF 111
Db      724 PCIHGACRDLGKGYKDCAPG-WSGTNCD-----INNNECESNPCVNGTGCKDM- 771

QY      112 DNSGMMC 118
Db      772 -TSGYVC 777

RESULT 4
US-10-072-012-471
; Sequence 471, Application US/10072012
; Publication No. US20040033493A1
; GENERAL INFORMATION:
; APPLICANT: Tchexnev, Velizar
; APPLICANT: Spytek, Kimberly
; APPLICANT: Zerhusen, Bryan
; APPLICANT: Patturajan, Meera
; APPLICANT: Shimkets, Richard
; APPLICANT: Li, Li
; APPLICANT: Gangolli, Esha
; APPLICANT: Padigaru, Muralidhara
; APPLICANT: Anderson, David W.
; APPLICANT: Rastelli, Luca
; APPLICANT: Miller, Charles E.
; APPLICANT: Gerlach, Valerie
; APPLICANT: Taupier Jr, Raymond J.
; APPLICANT: Gusev, Vladimir Y.
; APPLICANT: Colman, Steven D.
; APPLICANT: Wolenc, Adam R.
; APPLICANT: Pena, Carol E. A.
; APPLICANT: Furtak, Katarzyna
; APPLICANT: Grosse, William M.
; APPLICANT: Alsobrook II, John P.
; APPLICANT: Lepley, Denise M.
; APPLICANT: Rieger, Daniel K.
; APPLICANT: Burgess, Catherine E.
; TITLE OF INVENTION: Proteins and Nucleic Acids Encoding Same
; FILE REFERENCE: 21402-258
; CURRENT APPLICATION NUMBER: US/10/072,012
; CURRENT FILING DATE: 2002-01-31

```

```

; PRIOR APPLICATION NUMBER: 60/265,102
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: 60/265,514
; PRIOR FILING DATE: 2001-01-31
; PRIOR APPLICATION NUMBER: 60/265,517
; PRIOR FILING DATE: 2001-01-31
; PRIOR APPLICATION NUMBER: 60/265,412
; PRIOR FILING DATE: 2001-01-31
; PRIOR APPLICATION NUMBER: 60/265,395
; PRIOR FILING DATE: 2001-01-31
; PRIOR APPLICATION NUMBER: 60/266,406
; PRIOR FILING DATE: 2001-02-02
; PRIOR APPLICATION NUMBER: 60/266,767
; PRIOR FILING DATE: 2001-02-05
; PRIOR APPLICATION NUMBER: 60/267,057
; PRIOR FILING DATE: 2001-02-07
; PRIOR APPLICATION NUMBER: 60/266,975
; PRIOR FILING DATE: 2001-02-07
; PRIOR APPLICATION NUMBER: 60/267,459
; PRIOR FILING DATE: 2001-02-08
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 1391
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 471
; LENGTH: 2531
; TYPE: PRT
; ORGANISM: Rattus norvegicus
US-10-072-012-471

```

```

Query Match 12.7%; Score 95.5; DB 12; Length 2531;
Best Local Similarity 26.8%; Pred. No. 1.2;
Matches 34; Conservative 7; Mismatches 35; Indels 51; Gaps 7;

QY 15 ISACMLLCHGSLQHTFQOHLHHRPEGTCE-VIAAHC-----CNKN 56
Db 679 IDECAGSPCH-----NGGTCDGAGTCTCRCEGYHDPCTCLSEVNECNSN 723

QY 57 -----RIERSQTVKCSCLPGKVAGTTRNRPSCVDASIVIGKWCMEPCLEGECKTLP 111
Db 724 PCHGACRDLNGYKDCAPG-WSGTNCD-----INNNECESNPCVNGGTCKDM- 771

QY 112 DNSGMMC 118
Db 772 -TSGYVC 777

```

```

RESULT 5
US-10-190-115-29
; Sequence 29, Application US/10190115
; Publication No. US20030207394A1
; GENERAL INFORMATION:
; APPLICANT: Alsobrook, John P. II
; APPLICANT: Boldog, Ferenc L.
; APPLICANT: Burgess, Catherine E.
; APPLICANT: Casman, Stacie J.
; APPLICANT: Grosse, William M.
; APPLICANT: Gusev, Vladimir Y.
; APPLICANT: Ji, Weizhen
; APPLICANT: Lepley, Denise M.
; APPLICANT: Liu, Xiaohong
; APPLICANT: Mezzick, Amanda J.
; APPLICANT: Padigaru, Muralidhara
; APPLICANT: Patturajan, Meera
; APPLICANT: Rastelli, Luca
; APPLICANT: Shen, Lei
; APPLICANT: Shenoy, Suresh G.
; APPLICANT: Shimkets, Richard A.
; APPLICANT: Spaderna, Steven K.
; APPLICANT: Spytek, Kimberly A.
; APPLICANT: Szekeres, Edward S. Jr.
; APPLICANT: Taupier, Raymond J. Jr.
; APPLICANT: Tchernev, Velizar T.
; APPLICANT: Zerhusen, Bryan D.

```

```

; APPLICANT: Voss, Edward Z.
; TITLE OF INVENTION: NOVEL PROTEINS AND NUCLEIC ACIDS ENCODING SAME
; FILE REFERENCE: 21402-050 CIP
; CURRENT APPLICATION NUMBER: US/10/190,115
; CURRENT FILING DATE: 2003-02-10
; PRIOR APPLICATION NUMBER: 60/303,168
; PRIOR FILING DATE: 2001-07-05
; PRIOR APPLICATION NUMBER: 60/368,996
; PRIOR FILING DATE: 2002-04-01
; PRIOR APPLICATION NUMBER: 60/386,816
; PRIOR FILING DATE: 2002-06-07
; PRIOR APPLICATION NUMBER: 60/215,854
; PRIOR FILING DATE: 2000-07-03
; PRIOR APPLICATION NUMBER: 60/215,856
; PRIOR FILING DATE: 2000-07-03
; PRIOR APPLICATION NUMBER: 60/215,902
; PRIOR FILING DATE: 2000-07-03
; PRIOR APPLICATION NUMBER: 60/216,585,
; PRIOR FILING DATE: 2000-07-07
; PRIOR APPLICATION NUMBER: 60/216,586
; PRIOR FILING DATE: 2001-07-07
; PRIOR APPLICATION NUMBER: 60/216,722
; PRIOR FILING DATE: 2000-07-07
; PRIOR APPLICATION NUMBER: 60/218,622
; PRIOR FILING DATE: 2000-07-17
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 136
; SOFTWARE: Curaseqlist version 0.1
; SEQ ID NO 29
; LENGTH: 2531
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-190-115-29

```

```

Query Match 12.7%; Score 95.5; DB 15; Length 2531;
Best Local Similarity 26.8%; Pred. No. 1.2;
Matches 34; Conservative 7; Mismatches 35; Indels 51; Gaps 7;

QY 15 ISACMLLCHGSLQHTFQOHLHHRPEGTCE-VIAAHC-----CNKN 56
Db 679 IDECAGSPCH-----NGGTCDGAGTCTCRCEGYHDPCTCLSEVNECNSN 723

QY 57 -----RIERSQTVKCSCLPGKVAGTTRNRPSCVDASIVIGKWCMEPCLEGECKTLP 111
Db 724 PCHGACRDLNGYKDCAPG-WSGTNCD-----INNNECESNPCVNGGTCKDM- 771

QY 112 DNSGMMC 118
Db 772 -TSGYVC 777

```

```

RESULT 6
US-10-369-072-29
; Sequence 29, Application US/10369072
; Publication No. US20040014081A1
; GENERAL INFORMATION:
; APPLICANT: Alsobrook II, John P
; APPLICANT: Spaderna, Stephen K
; APPLICANT: Tchernev, Velizar
; APPLICANT: Liu, Xiaohong
; APPLICANT: Shenoy, Suresh
; APPLICANT: Spytek, Kimberly
; APPLICANT: Zerhusen, Bryan
; APPLICANT: Patturajan, Meera
; APPLICANT: Taupier, Raymond T
; APPLICANT: Rastelli, Luca
; APPLICANT: Grosse, William M
; APPLICANT: Szekeres, Edward S
; APPLICANT: Lepley, Denise M
; APPLICANT: Shen, Lei
; APPLICANT: Burgess, Catherine E
; APPLICANT: Shimkets, Richard
; APPLICANT: Padigaru, Muralidhara

```

; TITLE OF INVENTION: No. US20040014081a1el Proteins and Nucleic Acids Encoding Same
; FILE REFERENCE: 21402-050 CON2
; CURRENT APPLICATION NUMBER: US/10/369,072
; CURRENT FILING DATE: 2003-02-18
; PRIOR APPLICATION NUMBER: 10/174,372
; PRIOR FILING DATE: 2002-06-17
; PRIOR APPLICATION NUMBER: 08/898,994
; PRIOR FILING DATE: 2001-07-03
; PRIOR APPLICATION NUMBER: 60/215,854
; PRIOR FILING DATE: 2000-07-03
; PRIOR APPLICATION NUMBER: 60/215,856
; PRIOR FILING DATE: 2000-07-03
; PRIOR APPLICATION NUMBER: 60/215,902
; PRIOR FILING DATE: 2000-07-03
; PRIOR APPLICATION NUMBER: 60/216,585
; PRIOR FILING DATE: 2000-07-07
; PRIOR APPLICATION NUMBER: 60/216,586
; PRIOR FILING DATE: 2000-07-07
; PRIOR APPLICATION NUMBER: 60/216,722
; PRIOR FILING DATE: 2000-07-07
; PRIOR APPLICATION NUMBER: 60/218,622
; PRIOR FILING DATE: 2000-07-17
; PRIOR APPLICATION NUMBER: 60/218,992
; PRIOR FILING DATE: 2000-07-17
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 100
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 29
; LENGTH: 2531
; TYPE: PRT
; ORGANISM: Rattus norvegicus
US-10-369-072-29

Query Match 12.7%; Score 95.5; DB 15; Length 2531;
Best Local Similarity 26.8%; Pred. No. 1.2;
Matches 34; Conservative 7; Mismatches 35; Indels 51; Gaps 7;
QY 15 ISACMLLCHGSLQHTFQOHLHRRPEGGTCE-VIAAHR-----CNKN 56
DB 679 IDECAGSPCH-----NGGTCEDDGAGTCTCPCPEGYHPTCLSEVNECNSN 723
QY 57 -----RIEERSQTVKSCLPKGVAGTTRNRPSCVDASIVIGKWCMEPCLEGECKTL 111
DB 724 PCIHGACRDGLNGYKCDAPG-WSGTNCD-----INNNECESNPVNGGTCKDM- 771
QY 112 DNSGWC 118
DB 772 -TSGYVC 777

RESULT 7
US-09-216-393-110
; Sequence 110, Application US/09216393
; Patent No. US20010014447A1
; GENERAL INFORMATION:
; APPLICANT: Milhausen, Michael James
; TITLE OF INVENTION: TOXOPLASMA GONDII PROTEINS, NUCLEIC ACID MOLECULES, AND
; FILE REFERENCE: TX-1-C2
; CURRENT APPLICATION NUMBER: US/09/216,393
; CURRENT FILING DATE: 1998-12-18
; EARLIER APPLICATION NUMBER: 08/994,825
; EARLIER FILING DATE: 1997-12-19
; NUMBER OF SEQ ID NOS: 364
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 110
; LENGTH: 233
; TYPE: PRT
; ORGANISM: Toxoplasma gondii
US-09-216-393-110

Query Match 12.1%; Score 91.5; DB 9; Length 233;
Best Local Similarity 25.9%; Pred. No. 0.23;

Matches 28; Conservative 6; Mismatches 39; Indels 35; Gaps 4;
QY 18 CAMLLCHGSLQHTFQOHLHRRPEGGTCEVIAAHR-----CNKN-----RIEERS 62
DB 78 QCQDACHG-----GKTCEITKEHCCINDSDCNHGHGTCTNSNTNCCEAGF 122
QY 63 QTVKSCLPKGVAGTTRNRPSCVDASIVIGKWCMEPCLEGECKTL 110
DB 123 AGTNCSSSEKCSGKT-----CLSGHCNPATGACVCDPCHTGERCETL 165
RESULT 8
US-10-321-856-110
; Sequence 110, Application US/10321856
; Publication No. US20030194393A1
; GENERAL INFORMATION:
; APPLICANT: Milhausen, Michael James
; TITLE OF INVENTION: TOXOPLASMA GONDII PROTEINS, NUCLEIC ACID MOLECULES, AND USES THERE
; FILE REFERENCE: TX-1-C2-1
; CURRENT APPLICATION NUMBER: US/10/321,856
; CURRENT FILING DATE: 2002-12-17
; PRIOR APPLICATION NUMBER: 09/216,393
; PRIOR FILING DATE: 1998-12-18
; PRIOR APPLICATION NUMBER: 08/994,825
; PRIOR FILING DATE: 1997-12-19
; NUMBER OF SEQ ID NOS: 366
; SOFTWARE: Patentin version 3.1
; SEQ ID NO 110
; LENGTH: 233
; TYPE: PRT
; ORGANISM: Toxoplasma gondii
US-10-321-856-110

Query Match 12.1%; Score 91.5; DB 14; Length 233;
Best Local Similarity 25.9%; Pred. No. 0.23;
Matches 28; Conservative 6; Mismatches 39; Indels 35; Gaps 4;
QY 18 CAMLLCHGSLQHTFQOHLHRRPEGGTCEVIAAHR-----CNKN-----RIEERS 62
DB 78 QCQDACHG-----GKTCEITKEHCCINDSDCNHGHGTCTNSNTNCCEAGF 122
QY 63 QTVKSCLPKGVAGTTRNRPSCVDASIVIGKWCMEPCLEGECKTL 110
DB 123 AGTNCSSSEKCSGKT-----CLSGHCNPATGACVCDPCHTGERCETL 165

RESULT 9
US-10-190-115-28
; Sequence 28, Application US/10190115
; Publication No. US20030207394A1
; GENERAL INFORMATION:
; APPLICANT: Alsbrook, John P. II
; APPLICANT: Boldog, Ferenc L.
; APPLICANT: Burgess, Catherine E.
; APPLICANT: Casman, Stacie J.
; APPLICANT: Grosse, William M.
; APPLICANT: Gusev, Vladimir Y.
; APPLICANT: Ji, Weizhen
; APPLICANT: Lepley, Denise M.
; APPLICANT: Liu, Xiaohong
; APPLICANT: Mezick, Amanda J.
; APPLICANT: Padigaru, Muralidhara
; APPLICANT: Patturajan, Meera
; APPLICANT: Rastelli, Luca
; APPLICANT: Shen, Lei
; APPLICANT: Shenoy, Suresh G.
; APPLICANT: Shinkets, Richard A.
; APPLICANT: Spaderna, Steven K.
; APPLICANT: Sytek, Kimberly A.
; APPLICANT: Stekres, Edward S. Jr.
; APPLICANT: Taupier, Raymond J. Jr.
; APPLICANT: Tchernev, Velizar T.
; APPLICANT: Zerhusen, Bryan D.

```

; APPLICANT: Voss, Edward Z.
; TITLE OF INVENTION: NOVEL PROTEINS AND NUCLEIC ACIDS ENCODING SAME
; FILE REFERENCE: 21402-050 CIP
; CURRENT APPLICATION NUMBER: US/10/190,115
; CURRENT FILING DATE: 2003-02-10
; PRIOR APPLICATION NUMBER: 60/303,168
; PRIOR FILING DATE: 2001-07-05
; PRIOR APPLICATION NUMBER: 60/368,996
; PRIOR FILING DATE: 2002-04-01
; PRIOR APPLICATION NUMBER: 60/386,816
; PRIOR FILING DATE: 2002-06-07
; PRIOR APPLICATION NUMBER: 60/215,854
; PRIOR FILING DATE: 2000-07-03
; PRIOR APPLICATION NUMBER: 60/215,856
; PRIOR FILING DATE: 2000-07-03
; PRIOR APPLICATION NUMBER: 60/215,902
; PRIOR FILING DATE: 2000-07-03
; PRIOR APPLICATION NUMBER: 60/216,585
; PRIOR FILING DATE: 2000-07-07
; PRIOR APPLICATION NUMBER: 60/216,586
; PRIOR FILING DATE: 2000-07-07
; PRIOR APPLICATION NUMBER: 60/216,722
; PRIOR FILING DATE: 2000-07-03
; PRIOR APPLICATION NUMBER: 60/218,622
; PRIOR FILING DATE: 2000-07-07
; PRIOR APPLICATION NUMBER: 60/218,992
; PRIOR FILING DATE: 2000-07-17
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 136
; SOFTWARE: Curasequest version 0.1
; SEQ ID NO 28
; LENGTH: 2447
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-190-115-28

```

```

Query Match      11.7%; Score 88.5; DB 15; Length 2447;
Best Local Similarity 25.4%; Pred. No. 5.9;
Matches 30; Conservative 10; Mismatches 39; Indels 39; Gaps 5;

QY      41  GGTC-----VIAAHRCCNKNRIERSQTVKCSCLPGKVAGT-----TRNRPS 83
Db      404  GPRCEQDVNECASNPQNDGTCLDRIGDYSCICMPG-FGGTHCENELNECLSSPCLNRGK 462

QY      84  CVDAISIVIGKWCE-----MPCLEGECKTLPDMSGMCMCATGNK 123
Db      463  CLDQ----VSRFVCECPAGFGSEMCMQIDIDECSTPCLNGAKCIDLPNGYDCECAEGFK 517

```

```

RESULT 10
US-10-369-072-28
; Sequence 28, Application US/10369072
; Publication No. US20040014081A1
; GENERAL INFORMATION:
; APPLICANT: Alsobrook II, John P
; APPLICANT: Spaderna, Stephen K
; APPLICANT: Tchernev, Velizar
; APPLICANT: Liu, Xiaohong
; APPLICANT: Shenoy, Suresh
; APPLICANT: Spytek, Kimberly
; APPLICANT: Zerhusen, Bryan
; APPLICANT: Patturajan, Meera
; APPLICANT: Taupier, Raymond T
; APPLICANT: Rastelli, Luca
; APPLICANT: Grosse, William M
; APPLICANT: Szerkes, Edward S
; APPLICANT: Lepley, Denise M
; APPLICANT: Shen, Lei
; APPLICANT: Burgess, Catherine E
; APPLICANT: Shimkets, Richard
; APPLICANT: Padigaru, Muralidhara
; TITLE OF INVENTION: No. US20040014081A1el Proteins and Nucleic Acids Encoding Same
; FILE REFERENCE: 21402-050 CON2
; CURRENT APPLICATION NUMBER: US/10/369,072
; CURRENT FILING DATE: 2003-02-18

```

```

; PRIOR APPLICATION NUMBER: 10/174,372
; PRIOR FILING DATE: 2002-06-17
; PRIOR APPLICATION NUMBER: 09/898,994
; PRIOR FILING DATE: 2001-07-03
; PRIOR APPLICATION NUMBER: 60/215,854
; PRIOR FILING DATE: 2000-07-03
; PRIOR APPLICATION NUMBER: 60/215,856
; PRIOR FILING DATE: 2000-07-03
; PRIOR APPLICATION NUMBER: 60/215,902
; PRIOR FILING DATE: 2000-07-03
; PRIOR APPLICATION NUMBER: 60/216,585
; PRIOR FILING DATE: 2000-07-07
; PRIOR APPLICATION NUMBER: 60/216,586
; PRIOR FILING DATE: 2000-07-07
; PRIOR APPLICATION NUMBER: 60/216,722
; PRIOR FILING DATE: 2000-07-07
; PRIOR APPLICATION NUMBER: 60/218,622
; PRIOR FILING DATE: 2000-07-17
; PRIOR APPLICATION NUMBER: 60/218,992
; PRIOR FILING DATE: 2000-07-17
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 100
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 28
; LENGTH: 2447
; TYPE: PRT
; ORGANISM: Takifugu rubripes
US-10-369-072-28

Query Match      11.7%; Score 88.5; DB 15; Length 2447;
Best Local Similarity 25.4%; Pred. No. 5.9;
Matches 30; Conservative 10; Mismatches 39; Indels 39; Gaps 5;

QY      41  GGTC-----VIAAHRCCNKNRIERSQTVKCSCLPGKVAGT-----TRNRPS 83
Db      404  GPRCEQDVNECASNPQNDGTCLDRIGDYSCICMPG-FGGTHCENELNECLSSPCLNRGK 462

QY      84  CVDAISIVIGKWCE-----MPCLEGECKTLPDMSGMCMCATGNK 123
Db      463  CLDQ----VSRFVCECPAGFGSEMCMQIDIDECSTPCLNGAKCIDLPNGYDCECAEGFK 517

```

```

RESULT 11
US-09-764-853-773
; Sequence 773, Application US/09764853
; Patent No. US20020090672A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: P0206
; CURRENT APPLICATION NUMBER: US/09/764,853
; CURRENT FILING DATE: 2001-01-17
; Prior application data removed - consult PALM or file wrapper
; NUMBER OF SEQ ID NOS: 939
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 773
; LENGTH: 188
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (97)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (187)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (188)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
US-09-764-853-773

Query Match      11.5%; Score 87; DB 9; Length 188;
Best Local Similarity 25.0%; Pred. No. 0.53;

```

```

Matches 28; Conservative 9; Mismatches 45; Indels 30; Gaps 6;

QY 14 WISAC-----AMLLCHGSLQHTFQOHHLHRPEGGTCEVIAAHRCCNKNRIERSQTVKSC 69
Db 18 WQCTCDGCGGLFCDQDLNYC--THHSPCKNGATC-----SNSGQRSYT--CTC 62
QY 70 LPKVGAGTTRNRPSCVDASIVIGKWCCEMEPCLEGGCKTLPDNSGWMCATG 121
Db 63 RPYTG-----VCELELSE--CDSNPCRMGGCKQOEDGYXCLCPPG 103

RESULT 12
US-09-764-898-262
; Sequence 262, Application US/09764898
; Patent No. US2002090673A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PJ201
; CURRENT APPLICATION NUMBER: US/09/764,898
; CURRENT FILING DATE: 2001-01-17
; Prior application data removed - consult PALM or file wrapper
; NUMBER OF SEQ ID NOS: 311
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 262
; LENGTH: 188
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (97)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (187)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (188)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
US-09-764-898-262

Query Match 11.5%; Score 87; DB 9; Length 188;
Best Local Similarity 25.0%; Pred No, 0.53;
Matches 28; Conservative 9; Mismatches 45; Indels 30; Gaps 6;

QY 14 WISAC-----AMLLCHGSLQHTFQOHHLHRPEGGTCEVIAAHRCCNKNRIERSQTVKSC 69
Db 18 WQCTCDGCGGLFCDQDLNYC--THHSPCKNGATC-----SNSGQRSYT--CTC 62
QY 70 LPKVGAGTTRNRPSCVDASIVIGKWCCEMEPCLEGGCKTLPDNSGWMCATG 121
Db 63 RPYTG-----VCELELSE--CDSNPCRMGGCKQOEDGYXCLCPPG 103

RESULT 13
US-09-764-881-109
; Sequence 109, Application US/09764881
; Publication No. US20030125246A9
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PT207
; CURRENT APPLICATION NUMBER: US/09/764,881
; CURRENT FILING DATE: 2001-01-17
; Prior application data removed - refer to PALM or file wrapper
; NUMBER OF SEQ ID NOS: 192
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 109
; LENGTH: 188
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (97)

```



```
; CURRENT FILING DATE: 2002-09-13
; PRIOR APPLICATION NUMBER: 09/764,881
; PRIOR FILING DATE: 2001-01-17
; PRIOR APPLICATION NUMBER: 60/179,065
; PRIOR FILING DATE: 2000-01-31
; PRIOR APPLICATION NUMBER: 60/180,628
; PRIOR FILING DATE: 2000-02-04
; PRIOR APPLICATION NUMBER: 60/214,886
; PRIOR FILING DATE: 2000-06-28
; PRIOR APPLICATION NUMBER: 60/217,487
; PRIOR FILING DATE: 2000-07-11
; PRIOR APPLICATION NUMBER: 60/225,758
; PRIOR FILING DATE: 2000-08-14
; PRIOR APPLICATION NUMBER: 60/220,963
; PRIOR FILING DATE: 2000-07-26
; PRIOR APPLICATION NUMBER: 60/217,496
; PRIOR FILING DATE: 2000-07-11
; PRIOR APPLICATION NUMBER: 60/225,447
; PRIOR FILING DATE: 2000-08-14
; PRIOR APPLICATION NUMBER: 60/218,290
; PRIOR FILING DATE: 2000-07-14
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 192
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 109
; LENGTH: 188
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (97)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (187)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (188)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
US-10-242-747-109
```

```
Query Match 11.5%; Score 87; DB 15; Length 188;
Best Local Similarity 25.0%; Pred. No. 0.53;
Matches 28; Conservative 9; Mismatches 45; Indels 30; Gaps 6;

QY 14 WISAC-----AMLLCHGSLQHTFQOHLHRPEGGTCEVIAAHRCCNKNRIEBSOTVKCSC 69
Db | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
18 WQCTDEGWGGLFCDQDLYC--THSPCKNGATC-----SNSGORSYT--CTC 62

QY 70 LRGKVAGTTNRPSVDASIVIGKWCMEPCLEGECKTLPDNSGWMCATG 121
Db | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
63 RPYGTG-----VDCELELSE--CDSNFCRNGSGCKQDGDGYXCLCPPG 103
```

Search completed: April 8, 2004, 13:23:16
Job time : 40 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 8, 2004, 13:14:59 ; Search time 39 Seconds

(without alignments)
1075.998 Million cell updates/sec

Title: US-09-763-335-2

Perfect score: 754

Sequence: 1 MAMVSAMSWLYLWISACAM.....SGWMCATGNKIKTRIHPT 133

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1017041 seqs, 315518202 residues

Total number of hits satisfying chosen parameters: 1017041

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

- 1: sp_archaea:*
- 2: sp_bacteria:*
- 3: sp_fungi:*
- 4: sp_human:*
- 5: sp_invertebrate:*
- 6: sp_mammal:*
- 7: sp_mhc:*
- 8: sp_organelle:*
- 9: sp_phase:*
- 10: sp_plant:*
- 11: sp_rodent:*
- 12: sp_virus:*
- 13: sp_vertebrate:*
- 14: sp_unclassified:*
- 15: sp_rvirus:*
- 16: sp_bacteriaph:*
- 17: sp_archaeop:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	754	100.0	133	4	Q725A9
2	750	99.5	133	11	Q7TPG8
3	746	98.9	133	4	Q8TCL8
4	443	58.8	135	11	Q7TPG5
5	438	58.1	140	6	Q9NOD3
6	436.5	57.9	140	4	Q96LR4
7	435	57.7	140	6	Q95K92
8	432	57.3	131	4	Q8N3H0
9	429.5	57.0	132	11	Q7TPG6
10	429	56.9	131	11	Q7TPG7
11	415	55.0	126	11	Q8BV02
12	411.5	54.6	133	4	Q725A7
13	247.5	32.8	125	4	Q725A8
14	247.5	32.8	125	11	Q9IWE9
15	241	32.0	132	4	Q8IXR8
16	241	32.0	132	11	Q8C1V6

17	241	32.0	165	4	095902	095902 homo sapien
18	221.5	29.4	169	4	Q725A6	Q725A6 homo sapien
19	106	14.1	3871	5	Q20911	Q20911 caenorhabdi
20	95.5	12.7	2516	11	Q7TQ52	Q7TQ52 mus musculu
21	95.5	12.7	2526	11	Q7TQ51	Q7TQ51 mus musculu
22	95.5	12.7	2531	11	Q8K428	Q8K428 mus musculu
23	95.5	12.7	2531	11	Q7TQ50	Q7TQ50 mus musculu
24	94.5	12.5	2352	5	O61240	O61240 halocynthia
25	93	12.3	2531	5	O16004	O16004 lytechinus
26	92.5	12.3	2146	5	Q9VC97	Q9VC97 drosophila
27	89.5	11.9	2428	5	Q8I6X6	Q8I6X6 boophilus m
28	88.5	11.7	1212	13	O42347	O42347 gallus gall
29	88.5	11.7	2447	13	O13149	O13149 fugu rubrip
30	87	11.5	597	11	O35727	O35727 mus musculu
31	85	11.3	669	4	O75441	O75441 homo sapien
32	85	11.3	1511	4	O75412	O75412 homo sapien
33	85	11.3	1587	4	O00508	O00508 homo sapien
34	84.5	11.2	2528	13	O8AXP0	O8AXP0 cynops pyrr
35	84.5	11.2	2656	5	Q9GNU3	Q9GNU3 paracentrot
36	84	11.1	450	11	Q8KOH9	Q8KOH9 mus musculu
37	84	11.1	844	11	Q7TSG9	Q7TSG9 mus musculu
38	84	11.1	1405	11	Q8VHS2	Q8VHS2 mus musculu
39	83.5	11.1	729	13	Q7T3M4	Q7T3M4 brachydanio
40	83	11.0	305	6	Q9N028	Q9N028 macaca fasc
41	83	11.0	609	11	Q80YC5	Q80YC5 mus musculu
42	83	11.0	651	10	Q9FJE2	Q9FJE2 arabidopsis
43	83	11.0	674	5	Q8T4N9	Q8T4N9 strongyloce
44	83	11.0	1065	11	Q810H2	Q810H2 mus musculu
45	82.5	10.9	111	12	Q8JKE6	Q8JKE6 ectromelia

ALIGNMENTS

RESULT 1

Q725A9 PRELIMINARY; PRT; 133 AA.
AC Q725A9;
DT 01-OCT-2003 (TREMBLrel. 25, Created)
DT 01-OCT-2003 (TREMBLrel. 25, Last sequence update)
DT 01-OCT-2003 (TREMBLrel. 25, Last annotation update)
DE TAFAL.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Tang Y.T., Emtage P., Funk W., Hu T., Arterburn M., Park E., Rupp F.;
RT "TAFAL: A Novel Secreted Family with Homology to CC-chemokines.";
RL Submitted (JUN-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY325114; AAP92406.1; -.
SQ SEQUENCE 133 AA; 14901 MW; C5DD1C7E5F997386 CRC64;

Query Match	100.0%	Score 754;	DB 4;	Length 133;
Best Local Similarity	100.0%;	Pred. No. 2.6e-80;		
Matches 133;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
QY	1	MAMVSAMSWLYLWISACAMLLCHGSLQHTFQOHLHRLPEPGTCEVIAAHRCCNKNRIE	60	
Db	1	MAMVSAMSWLYLWISACAMLLCHGSLQHTFQOHLHRLPEPGTCEVIAAHRCCNKNRIE	60	
QY	61	RSQTVKCSCLPKVAGTTRNRPSVCVDASIVIGKWCEMPCLEGBECKTLPNSGWMCAT	120	
Db	61	RSQTVKCSCLPKVAGTTRNRPSVCVDASIVIGKWCEMPCLEGBECKTLPNSGWMCAT	120	
QY	121	GNKIKTRIHPT 133		
Db	121	GNKIKTRIHPT 133		

RESULT 2

Q7TPG8

```

ID Q7PG8 PRELIMINARY; PRT; 133 AA.
AC Q7PG8;
DT 01-OCT-2003 (T-EMBLrel. 25, Created)
DT 01-OCT-2003 (T-EMBLrel. 25, Last sequence update)
DT 01-OCT-2003 (T-EMBLrel. 25, Last annotation update)
DE TAF1.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J;
RA Tang Y.T., Emtage P., Funk W., Hu T., Arterburn M., Park E., Rupp F.;
RT "TAPA: A Novel Secreted Family with Homology to CC-chemokines.";
RL Submitted (JUN-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY325120; AAP92412.1; -.
SQ SEQUENCE 133 AA; 14873 MW; C5C20C764F8E7386 CRC64;

Query Match 99.5%; Score 750; DB 11; Length 133;
Best Local Similarity 99.2%; Pred. No. 7.7e-80;
Matches 132; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 MAMVSAMSWLYLWISACAMLLCHGSLQHTFQQHHLRPEGGTCEVIAAHRCCNKNRIE 60
Dd 1 MAMVSAMSWLYLWISACAMLLCHGSLQHTFQQHHLRPEGGTCEVIAAHRCCNKNRIE 60
QY 61 RSQTVKSCSLPGKVGATRRPSCVDASIVIGKWCMEPCLEGECKTLPDNSGWCAT 120
Dd 61 RSQTVKSCSLPGKVGATRRPSCVDASIVIGKWCMEPCLEGECKTLPDNSGWCAT 120
QY 121 GNKIKTRIHPT 133
Dd 121 GNKIKTRIHPT 133

RESULT 3
Q8TCL8 PRELIMINARY; PRT; 133 AA.
AC Q8TCL8;
DT 01-JUN-2002 (T-EMBLrel. 21, Created)
DT 01-JUN-2002 (T-EMBLrel. 21, Last sequence update)
DT 01-OCT-2002 (T-EMBLrel. 22, Last annotation update)
DE Hypothetical protein.
GN DXFP566B064.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Kidney;
RA Bloeker H., Boecher M., Brandt P., Mewes H.W., Gassenhuber J.,
RA Wiemann S.;
RL Submitted (MAR-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AL713702; CAD28501.1; -.
KW Hypothetical protein.
SQ SEQUENCE 133 AA; 15030 MW; C5C0802F0F997386 CRC64;

Query Match 98.9%; Score 746; DB 4; Length 133;
Best Local Similarity 99.2%; Pred. No. 2.3e-79;
Matches 132; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 MAMVSAMSWLYLWISACAMLLCHGSLQHTFQQHHLRPEGGTCEVIAAHRCCNKNRIE 60
Dd 1 MAMVSAMSWLYLWISACAMLLCHGSLQHTFQQHHLRPEGGTCEVIAAHRCCNKNRIE 60
QY 61 RSQTVKSCSLPGKVGATRRPSCVDASIVIGKWCMEPCLEGECKTLPDNSGWCAT 120
Dd 61 RSQTVKSCSLPGKVGATRRPSCVDASIVIGKWCMEPCLEGECKTLPDNSGWCAT 120
QY 121 GNKIKTRIHPT 133
Dd 121 GNKIKTRIHPT 133

```

```

Db 121 GNKIKTRIHPT 133

RESULT 4
Q7TPG5 PRELIMINARY; PRT; 135 AA.
ID Q7TPG5;
AC Q7TPG5;
DT 01-OCT-2003 (T-EMBLrel. 25, Created)
DT 01-OCT-2003 (T-EMBLrel. 25, Last sequence update)
DT 01-OCT-2003 (T-EMBLrel. 25, Last annotation update)
DE TAF4.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J;
RA Tang Y.T., Emtage P., Funk W., Hu T., Arterburn M., Park E., Rupp F.;
RT "TAPA: A Novel Secreted Family with Homology to CC-chemokines.";
RL Submitted (JUN-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY325123; AAP92415.1; -.
SQ SEQUENCE 135 AA; 15018 MW; 38C83231992D3907 CRC64;

Query Match 58.8%; Score 443; DB 11; Length 135;
Best Local Similarity 62.6%; Pred. No. 6.7e-44;
Matches 77; Conservative 17; Mismatches 25; Indels 4; Gaps 2;

QY 9 WVLVLMISACAMLLCHGSLQHT--FOQHLRPEGGTCEVIAAHRCCNKNRIERSQTVK 66
Dd 13 WVLVLMISACAMLLCHGSLQHT--FOQHLRPEGGTCEVIAAHRCCNKNRIERSQTVK 70
QY 67 CSCLPGKVGATRRPSCVDASIVIGKWCMEPCLEGECKTLPDNSGWCATGNKIKT 126
Dd 71 CSCLPGKVGATRRPSCVDASIVIGKWCMEPCLEGECKTLPDNSGWCATGNKIKT 130
QY 127 TRI 129
Dd 131 TKV 133

RESULT 5
Q9NOD3 PRELIMINARY; PRT; 140 AA.
ID Q9NOD3;
AC Q9NOD3;
DT 01-OCT-2000 (T-EMBLrel. 15, Created)
DT 01-OCT-2000 (T-EMBLrel. 15, Last sequence update)
DT 01-OCT-2000 (T-EMBLrel. 15, Last annotation update)
DE Unnamed protein product.
OS Macaca fascicularis (Crab eating macaque) (Cynomolgus monkey).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Cercopitheidae;
OC Cercopitheidae; Macaca.
OX NCBI_TaxID=9541;
RN [1]
RP SEQUENCE FROM N.A.
RA Osada N., Hida M., Kusuda J., Tanuma R., Iseki K., Hirai M., Terao K.,
RA Suzuki Y., Sugano S., Hashimoto K.;
RT "Isolation of full-length cDNA clones from macaque brain cDNA
RT libraries.";
RL Submitted (JUL-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AB045997; BAB01579.1; -.
SQ SEQUENCE 140 AA; 15640 MW; 173DA66AB6C97F03 CRC64;

Query Match 58.1%; Score 438; DB 6; Length 140;
Best Local Similarity 62.8%; Pred. No. 2.7e-43;
Matches 76; Conservative 14; Mismatches 31; Indels 0; Gaps 0;

QY 9 WVLVLMISACAMLLCHGSLQHTFQQHHLRPEGGTCEVIAAHRCCNKNRIERSQTVKCS 68
Dd 18 WVLVLMISACAMLLCHGSLQHTFQQHHLRPEGGTCEVIAAHRCCNKNRIERSQTVKCS 77
QY 69 CLPGKVGATRRPSCVDASIVIGKWCMEPCLEGECKTLPDNSGWCATGNKIKTTR 128

```

```

Db      78 CFFGQVAGTTAQPSCVEASIVIQKWCNNPCLEGECKVLDPYSGWSSGNNKVKTK 137
QY      129 I 129
Db      138 V 138

RESULT 6
Q96LR4 ID Q96LR4 PRELIMINARY; PRT; 140 AA.
AC Q96LR4;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Hypothetical protein FLJ25161 (Similar to expressed sequence AW049604)
DE (TAF4A).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Brain;
RA Tashiro H., Yamazaki M., Watanabe K., Kumagai A., Itakura S.,
RA Fukuzumi Y., Fujimori Y., Komiyama M., Suzuki Y., Hata H.,
RA Nakagawa K., Mizuno S., Morinaga M., Kawamura M., Sugiyama T.,
RA Irie R., Otsubi T., Sato H., Nishikawa T., Sugiyama A., Kawakami B.,
RA Nagai K., Isogai T., Sugano S.;
RA "NEDO human cDNA sequencing project.";
RL Submitted (OCT-2001) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Brain;
RA Strausberg R.;
RN [3]
RP SEQUENCE FROM N.A.
RA Tang Y.T., Emtage P., Funk W., Hu T., Arterburn M., Park E., Rupp F.;
RL "TAPA: A Novel Secreted Family with Homology to CC-chemokines.";
RT Submitted (JUN-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; AK057890; BAB71606.1; -
DR EMBL; BC031566; AAH31566.1; -
DR EMBL; AY325117; AAP92409.1; -
KW Hypothetical protein.
SQ SEQUENCE 140 AA; 15682 MW; 173254FB8B526009 CRC64;

Query Match 57.9%; Score 436.5; DB 4; Length 140;
Best Local Similarity 61.1%; Pred. No. 4e-43;
Matches 80; Conservative 16; Mismatches 30; Indels 5; Gaps 2;

QY 4 VSAMSVLYLWISACAMLLCHGSIQHTFQOH---HL--HRPEGTCEVIAAHRCCNNRI 58
Db 8 VCAKSVLLSWFLAYVLMVCCKLMSASSQHLRGHAGHQIKQGTCEVAVHRCNNRI 67

QY 59 EERSQIVKSCFLPGKVGATTRNRPSCVDASIVIGKWCNMEPCLEGECKTLPDINSWMC 118
Db 68 EERSQIVKSCFFGQVAGTTAQPSCVEASIVIQKWCNNPCLEGECKVLDPYSGWSC 127

QY 119 ATGNKIKTKTR 129
Db 128 SSGNKVKTKV 138

RESULT 7
Q95K92 ID Q95K92 PRELIMINARY; PRT; 140 AA.
AC Q95K92;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Hypothetical protein.
OS Macaca fascicularis (Crab eating macaque) (Cynomolgus monkey).

```

```

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Cercopitheciidae;
OC Cercopitheciinae; Macaca.
OX NCBI_TaxID=9541;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Medulla oblongata;
RA Osada N., Hida M., Kusuda J., Tanuma R., Iseki K., Hirai M., Terao K.,
RA Suzuki Y., Sugano S., Hashimoto K.;
RT "Isolation of full-length cDNA clones from macaque brain cDNA
RT libraries.";
RL Submitted (JUN-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AB063062; BAB60784.1; -
KW Hypothetical protein.
SQ SEQUENCE 140 AA; 15654 MW; 17324670B6C97F19 CRC64;

Query Match 57.7%; Score 435; DB 6; Length 140;
Best Local Similarity 62.0%; Pred. No. 6.1e-43;
Matches 75; Conservative 15; Mismatches 31; Indels 0; Gaps 0;

QY 9 WLYLWISACAMLLCHGSIQHTFQOHHLHRPEGTCEVIAAHRCCNNRIEERSQIVKCS 68
Db 18 WLFAYLWVCCKLMSASSQHLRGHAGHQIKQGTCEVAVHRCNNRIEERSQIVKCS 77

QY 69 CLPGKVGATTRNRPSCVDASIVIGKWCNMEPCLEGECKTLPDINSWMCATGNKIKTKR 128
Db 78 CFFGQVAGTTAQPSCVEASIVIQKWCNNPCLEGECKVLDPYSGWSSGNNKVKTK 137

QY 129 I 129
Db 138 V 138

RESULT 8
Q8N3H0 ID Q8N3H0 PRELIMINARY; PRT; 131 AA.
AC Q8N3H0;
DT 01-OCT-2002 (TrEMBLrel. 22, Created)
DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Hypothetical protein (TAPA2).
GN DKFZF761E1217.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Amygdala;
RA Ansoorge W., Witkner U., Mewes H.W., Weil B., Wiemann S.;
RL Submitted (JUL-2002) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RA Tang Y.T., Emtage P., Funk W., Hu T., Arterburn M., Park E., Rupp F.;
RL "TAPA: A Novel Secreted Family with Homology to CC-chemokines.";
RT Submitted (JUN-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; AL834160; CAD38865.1; -
DR EMBL; AY325115; AAP92407.1; -
KW Hypothetical protein.
SQ SEQUENCE 131 AA; 14620 MW; 6D7E0DB2E59965E4 CRC64;

Query Match 57.3%; Score 432; DB 4; Length 131;
Best Local Similarity 61.2%; Pred. No. 1.3e-42;
Matches 74; Conservative 14; Mismatches 23; Indels 10; Gaps 1;

QY 9 WLYLWISACAMLLCHGSIQHTFQOHHLHRPEGTCEVIAAHRCCNNRIEERSQIVKCS 68
Db 19 FIVTLW-----GKVSSANHHKAAHVKTGTCEVVALHRCNNRIEERSQIVKCS 68

QY 69 CLPGKVGATTRNRPSCVDASIVIGKWCNMEPCLEGECKTLPDINSWMCATGNKIKTKR 128
Db 69 CFFGQVAGTTAQPSCVDASIVIQKWCNNPCLEGECKVLDPYSGWSSGNNKVKTKR 128

```

```

QY      129 I 129
Db      129 V 129

RESULT 9
Q7TPG6
ID      Q7TPG6      PRELIMINARY;      PRT;      132 AA.
AC      Q7TPG6;
DT      01-OCT-2003 (TREMBlrel. 25, Created)
DT      01-OCT-2003 (TREMBlrel. 25, Last sequence update)
DT      01-OCT-2003 (TREMBlrel. 25, Last annotation update)
DE      TAPAF3.
OS      Mus musculus (Mouse).
OC      Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC      Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX      NCBI_TaxID=10090;
RN      [1]
RP      SEQUENCE FROM N.A.
RC      STRAIN=C57BL/6J;
RA      Tang Y.T., Emtage P., Funk W., Hu T., Arterburn M., Park E., Rupp F.;
RT      "TAPA: A Novel Secreted Family with Homology to CC-Chemokines.";
RL      Submitted (JUN-2003) to the EMBL/GenBank/DBJ databases.
DR      EMBL; AY325122; AAP92414.1; -.
SQ      SEQUENCE 132 AA; 14426 MW; 2D60C45AE1BA3F00 CRC64;

Query Match      57.0%; Score 429.5; DB 11; Length 132;
Best Local Similarity 61.5%; Pred. No. 2.5e-42;
Matches 83; Conservative 13; Mismatches 16; Indels 23; Gaps 4;

QY      5 SAMSWSLYL---WISACAMLLCHGSLQHTFQOHLHRP-----EGGTCVIAAHRCCN 54
Db      9 SAGSWVLALCLAWLWCP-----ASASLQ-----PPTSAILVKGTCEVIAAHRCCN 55

QY      55 KNRIERSQTVKSCULPGKVAGTTRNRPSCVDASIVIGKWCMEPCLEGECKTLPDNS 114
Db      56 RNRIERSQTVKSCLSGQVAGTTRAKPSCVDASIVLQKWCQMEPCLLGECKVLPDLS 115

QY      115 GWMCATGNKIKTTRI 129
Db      116 GWSGSSGHKVKTKV 130

RESULT 10
Q7TPG7
ID      Q7TPG7      PRELIMINARY;      PRT;      131 AA.
AC      Q7TPG7;
DT      01-OCT-2003 (TREMBlrel. 25, Created)
DT      01-OCT-2003 (TREMBlrel. 25, Last sequence update)
DT      01-OCT-2003 (TREMBlrel. 25, Last annotation update)
DE      TAPAF2.
OS      Mus musculus (Mouse).
OC      Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC      Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX      NCBI_TaxID=10090;
RN      [1]
RP      SEQUENCE FROM N.A.
RC      STRAIN=C57BL/6J;
RA      Tang Y.T., Emtage P., Funk W., Hu T., Arterburn M., Park E., Rupp F.;
RT      "TAPA: A Novel Secreted Family with Homology to CC-Chemokines.";
RL      Submitted (JUN-2003) to the EMBL/GenBank/DBJ databases.
DR      EMBL; AY325121; AAP92413.1; -.
SQ      SEQUENCE 131 AA; 14647 MW; FE70D61916A95F47 CRC64;

Query Match      56.9%; Score 429; DB 11; Length 131;
Best Local Similarity 61.2%; Pred. No. 2.8e-42;
Matches 74; Conservative 12; Mismatches 25; Indels 10; Gaps 1;

QY      9 WVLWLSACAMLLCHGSLQHTFQOHLHRPEGGTCVIAAHRCCNRIEERSQTVKCS 68
Db      19 FIVTLW-----GKAVSSANHHKAHVRTGCEVIAAHRCCNKNRIEERSQTVKCS 69

QY      69 CLPGKVAGTTRNRPSCVDASIVIGKWCMEPCLEGECKTLPDNSGWMCATGNKIKTTR 128

```

```

Db      69 CFPQGVAGTTRAPSCVDASIVIGKWCMEPCLEGECKVLPDRKWSGSSGNKVKTTR 128
QY      129 I 129
Db      129 V 129

RESULT 11
Q8BV02
ID      Q8BV02      PRELIMINARY;      PRT;      126 AA.
AC      Q8BV02;
DT      01-MAR-2003 (TREMBlrel. 23, Created)
DT      01-MAR-2003 (TREMBlrel. 23, Last sequence update)
DT      01-JUN-2003 (TREMBlrel. 24, Last annotation update)
DE      CDNA FLJ25161 FIS.
GN      C130034I18RIK.
OS      Mus musculus (Mouse).
OC      Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC      Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX      NCBI_TaxID=10090;
RN      [1]
RP      SEQUENCE FROM N.A.
RC      STRAIN=C57BL/6J; TISSUE=Head;
RX      MEDLINE=22354683; PubMed=12466851;
RA      The FANTOM Consortium,
RT      the RIKEN Genome Exploration Research Group Phase I & II Team;
RL      "Analysis of the mouse transcriptome based on functional annotation of
RL      60,770 full-length cDNAs.";
DR      EMBL; AK081531; BAC38247.1; -.
DR      MGD; MGI:2444563; C130034I18RIK.
SQ      SEQUENCE 126 AA; 13975 MW; D2AA1F362D44ACBA CRC64;

Query Match      55.0%; Score 415; DB 11; Length 126;
Best Local Similarity 62.9%; Pred. No. 1.2e-40;
Matches 73; Conservative 14; Mismatches 25; Indels 4; Gaps 2;

QY      9 WVLWLSACAMLLCHGSLQHT---FQOHLHRPEGGTCVIAAHRCCNRIEERSQTVK 66
Db      13 WLVLTYVLMVCKKMSASSQHLRGHAGHHLKP--GTCEVAVHRCNKNRIEERSQTVK 70

QY      67 CSCLPGKVAGTTRNRPSCVDASIVIGKWCMEPCLEGECKTLPDNSGWMCATGN 122
Db      71 CSCFFQGVAGTTRAPSCVVEAIVIEKWNPCLEGEDECKVLPDSSGWSGSSGN 126

RESULT 12
Q7Z5A8
ID      Q7Z5A8      PRELIMINARY;      PRT;      133 AA.
AC      Q7Z5A8;
DT      01-OCT-2003 (TREMBlrel. 25, Created)
DT      01-OCT-2003 (TREMBlrel. 25, Last sequence update)
DT      01-OCT-2003 (TREMBlrel. 25, Last annotation update)
DE      TAPAF3.
OS      Homo sapiens (Human).
OC      Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC      Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX      NCBI_TaxID=9606;
RN      [1]
RP      SEQUENCE FROM N.A.
RC      Tang Y.T., Emtage P., Funk W., Hu T., Arterburn M., Park E., Rupp F.;
RT      "TAPA: A Novel Secreted Family with Homology to CC-Chemokines.";
RL      Submitted (JUN-2003) to the EMBL/GenBank/DBJ databases.
DR      EMBL; AY325116; AAP92408.1; -.
SQ      SEQUENCE 133 AA; 14776 MW; C074C703817D008D CRC64;

Query Match      54.6%; Score 411.5; DB 4; Length 133;
Best Local Similarity 58.5%; Pred. No. 3.2e-40;
Matches 76; Conservative 15; Mismatches 26; Indels 13; Gaps 2;

QY      5 SAMSWSYL-----YLWTSACAMLLCHGSLQHTFQOHLHRPEGGTCVIAAHRCCNRIE 59

```

Db 10 STGWLALCLANLWHLTLAALOPPTATVLVQ-----GTCEVIAAHRCCNKNRIE 61
QY 60 ERSQTVKSCULPGKVGAGTRNRPSVDASIVIGKWCMEPCLEGECKTLDPNSGMCA 119
Db 62 ERSQTVKSCFSGGVAGTRAKPSCVDASIVLQRMWCMEPCLEGECKTLDPNSGMCA 121
QY 120 TGNKIKTKTRI 129
Db 122 SGHKVKTKV 131

RESULT 13
Q7Z5A7 PRELIMINARY; PRT; 125 AA.
AC Q7Z5A7
DT 01-OCT-2003 (TREMBlrel. 25, Created)
DT 01-OCT-2003 (TREMBlrel. 25, Last sequence update)
DT 01-OCT-2003 (TREMBlrel. 25, Last annotation update)
DE TAFAS.
OS Homo sapiens (Human)
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Tang Y. T., Entage P., Funk W., Hu T., Arterburn M., Park E., Rupp F.;
RT "TAFAS: A Novel Secreted Family with Homology to CC-chemokines";
RL Submitted (JUN-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY325118; AAP92410.1; -;
SQ SEQUENCE 125 AA; 13592 MW; A8BCA8D09F86404A CRC64;

Query Match 32.8%; Score 247.5; DB 4; Length 125;
Best Local Similarity 44.3%; Pred. No. 4.7e-21;
Matches 54; Conservative 14; Mismatches 45; Indels 9; Gaps 3;
QY 13 LWSACAMLLCHGSLQHTFOQHLHRPEG-----VHAQFLKEGQLAGTCEIVLDRDSSQPRRTIARQTA 68
Db 7 LWSACAMLLCHGSLQHTFOQHLHRPEG-----VHAQFLKEGQLAGTCEIVLDRDSSQPRRTIARQTA 62
QY 69 CLPGKVGAGTRNRPSVDASIVIGKWCMEPCLEGECKTLDPNSGMCA-TGNKIKTT 127
Db 63 CRKQIAGTTRARPACVDARIITKQWCDMLPCLEGGCDLLINRSWGTCTQPGRIKT 122
QY 128 RI 129
Db 123 TV 124

RESULT 14
Q91WE9 PRELIMINARY; PRT; 125 AA.
AC Q91WE9
DT 01-DEC-2001 (TREMBlrel. 19, Created)
DT 01-DEC-2001 (TREMBlrel. 19, Last sequence update)
DT 01-OCT-2003 (TREMBlrel. 25, Last annotation update)
DE Hypothetical protein (TAFAS).
GN AW049604
OS Mus musculus (Mouse)
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Kidney;
RA Strausberg R.;
RL Submitted (OCT-2001) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J;
RA Tang Y. T., Entage P., Funk W., Hu T., Arterburn M., Park E., Rupp F.;
RT "TAFAS: A Novel Secreted Family with Homology to CC-chemokines";
RL Submitted (JUN-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC015306; AAH15306.1; -;

DR EMBL; AY325124; AAP92416.1; -;
DR MGD; MGI:2146182; AW049604.
KW Hypothetical protein.
SQ SEQUENCE 125 AA; 13592 MW; A8BCA8D09F86404A CRC64;

Query Match 32.8%; Score 247.5; DB 11; Length 125;
Best Local Similarity 44.3%; Pred. No. 4.7e-21;
Matches 54; Conservative 14; Mismatches 45; Indels 9; Gaps 3;
QY 13 LWSACAMLLCHGSLQHTFOQHLHRPEG-----GTCEVIAAHRCCNKNRIEERSQTVKCS 68
Db 7 LWSACAMLLCHGSLQHTFOQHLHRPEG-----VHAQFLKEGQLAGTCEIVLDRDSSQPRRTIARQTA 62
QY 69 CLPGKVGAGTRNRPSVDASIVIGKWCMEPCLEGECKTLDPNSGMCA-TGNKIKTT 127
Db 63 CRKQIAGTTRARPACVDARIITKQWCDMLPCLEGGCDLLINRSWGTCTQPGRIKT 122
QY 128 RI 129
Db 123 TV 124

RESULT 15
Q81XR8 PRELIMINARY; PRT; 132 AA.
AC Q81XR8
DT 01-MAR-2003 (TREMBlrel. 23, Created)
DT 01-MAR-2003 (TREMBlrel. 23, Last sequence update)
DT 01-MAR-2003 (TREMBlrel. 23, Last annotation update)
DE Similar to expressed sequence AW049604.
OS Homo sapiens (Human)
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Brain;
RA Strausberg R.;
RL Submitted (NOV-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC039396; AAH39396.1; -;
SQ SEQUENCE 132 AA; 14300 MW; 16BFE9F574FBB066 CRC64;

Query Match 32.0%; Score 241; DB 4; Length 132;
Best Local Similarity 38.4%; Pred. No. 2.9e-20;
Matches 48; Conservative 23; Mismatches 42; Indels 12; Gaps 2;
QY 6 AMSWLYLWISACAMLLCHGSLQHTFOQHLHRPEGTCVIAAHRCCNKNRIEERSQTV 65
Db 18 SMSSTWAFWILASLLIAYCS-----QLAAGTCEIVLDRDSSQPRRTIARQTA 66
QY 66 KCSCLPGKVGAGTRNRPSVDASIVIGKWCMEPCLEGECKTLDPNSGMCA-TGNKI 124
Db 67 RCACRKGQIAGTTRARPACVDARIITKQWCDMLPCLEGGCDLLINRSWGTCTQPGRI 126
QY 125 KITRI 129
Db 127 KITTV 131

Search completed: April 8, 2004, 13:17:32
Job time : 40 secs